

De invloed van opbrengstindicatoren op het functioneren van scholen in het primair onderwijs

Lex Borghans en Trudie Schils

Universiteit Maastricht

Met medewerking van

Annemarie van Langen en Bianca Leest

ITS Nijmegen

Eind rapportage oktober 2015

Inhoudsopgave

INHOUDSOPGAVE	2
1 INTRODUCTIE	3
2 MOGELIJKE VORMEN VAN STRATEGISCH GEDRAG	5
2.1 ALGEMENE THEORIE OVER HET GEBRUIK VAN KWALITEITSINDICATOREN	5
2.2 LITERATUUR OVER HET GEBRUIK VAN KWALITEITSINDICATOREN	10
2.3 THEORIE OVER STRATEGISCH GEDRAG	12
2.4 STRATEGISCH GEDRAG IN DE LITERATUUR	14
2.5 OVERZICHT TOEZICHTKADER VAN DE INSPECTIE	17
2.6 OVERZICHT BEKOSTIGING VAN SCHOLEN	18
2.7 HOE KUN JE STRATEGISCH GEDRAG METEN?	19
3 VERKENNING STRATEGISCH GEDRAG IN HET PRIMAIR ONDERWIJS	22
3.1 GESPREKKEN MET SCHOOLLEIDERS EN LEERKRACHTEN	22
3.2 GESPREK MET DE INSPECTIE	25
3.3 KEUZE VAN THEMA'S	25
4 ANALYSES VAN SCHOOLPRAKTIJK OP BASIS VAN VRAGENLIJSTEN	27
4.1 SCHOOLLEIDERS EN LEERKRACHTEN.....	27
4.2 OUDERS	43
5 ANALYSE VAN TWEE OPBRENGSTINDICATOREN OP BASIS VAN MICRODATA	49
5.1 VERTRAAGDE DOORSTROOM.....	49
5.2 DE EINDTOETS BASISONDERWIJS	53
6 CONCLUSIES EN AANBEVELINGEN	61
BIJLAGE A KERNKADER PRIMAIR ONDERWIJS	66
BIJLAGE B GESPREKSLEIDRAAD VOOR GESPREKKEN MET SCHOOLLEIDERS EN LEERKRACHTEN	68
BIJLAGE C VRAGENLIJST OUDERS	70
BIJLAGE D VRAGENLIJST SCHOOLLEIDERS EN LEERKRACHTEN	72

1 Introductie

De Inspectie van het Onderwijs (inspectie) houdt met behulp van observaties en indicatoren toezicht op de scholen. In het kader van het risicogericht toezicht wordt daarbij allereerst bekeken of er sprake is van verhoogd risico om vervolgens – als dat zo is – in een kwaliteitsonderzoek te bepalen of een school voldoet aan de kwaliteitseisen. De procedures die hierbij gehanteerd worden, zijn vastgelegd in het toezichtkader. Scholen weten dus hoe ze door de inspectie worden beoordeeld. Als gevolg hiervan passen scholen hun beleid aan. Dat is ook de bedoeling van het toezicht. De inspectie heeft immers als doel dat scholen voldoen aan de kwaliteitseisen. Als die kwaliteit zonder het toezicht niet wordt geleverd, kan er dus een positieve stimulerende werking uitgaan van de beoordeling. Het toezicht kan echter ook ongewenste effecten hebben. Beoordelingen op basis van observaties en indicatoren, hebben in hun aard een globaal karakter. Dat betekent dat er ruimte kan bestaan voor scholen om weliswaar te voldoen aan de beoordeling op basis van de indicatoren, zonder daarbij daadwerkelijk het onderwijs te verbeteren. Ook kunnen er situaties ontstaan waarin scholen die beleid voeren dat in de praktijk tot kwalitatief voldoende onderwijs leidt, worden beperkt door de eisen die het toezichtkader hen oplegt. Daarnaast kunnen scholen geconfronteerd worden met onzekerheid bij hun beoordeling.

In dit rapport onderzoeken we in hoeverre er in het Nederlandse basisonderwijs sprake lijkt te zijn van dergelijk ongewenst strategisch gedrag, en bespreken we wat de inspectie kan doen om dergelijk strategisch gedrag te monitoren en de kans hier op te verkleinen. Dat doen we vanuit een gedragseconomisch perspectief.

Eerst verkennen we de mogelijkheden voor ongewenst strategisch gedrag. Dit doen we op basis van de gedragseconomische theorie en de mogelijkheden voor strategisch gedrag, die in het huidige toezichtkader zitten. Daarbij is het van belang om niet alleen het toezichtkader zelf, maar ook andere regelgeving in beschouwing te nemen die invloed kan hebben op de strategie van scholen. Hierbij gaat het met name om de bekostiging van scholen. Bij het vaststellen van eventueel strategisch gedrag doen zich twee problemen voor.

Ten eerste komt strategisch gedrag voort uit het feit dat het feitelijke gedrag van scholen onvolledig tot uiting komt in de gegevens waarover de inspectie beschikt. Strategisch gedrag kan daardoor niet op een eenvoudige wijze uit deze gegevens worden afgeleid. Daarnaast is het voor de beoordeling van strategisch gedrag nodig om te weten wat het gewenste gedrag van scholen zou zijn geweest. Ook dit is niet eenvoudig op basis van de beschikbare gegevens vast te stellen. Op basis van de literatuur bespreken we wat de mogelijkheden zijn om strategisch gedrag vast te stellen en te beoordelen. Ook kijken we naar de literatuur omdat in het buitenland ervaring is opgedaan met zowel het werken met kwaliteitsindicatoren, als met strategische gedrag van scholen in reactie op bekostigings- en beoordelingsregels van de overheid. Deze stap wordt uitgewerkt in hoofdstuk 2 van dit rapport.

Daarnaast zijn er verkennende gesprekken gevoerd met de inspectie, schoolleiders en leerkrachten van basisscholen, de Algemene Vereniging van Schoolleiders en de PO-Raad, om het beeld van de invloed dat het toezichtkader, en met name de opbrengstindicatoren als onderdeel daarvan, op het gedrag van scholen kan hebben verder aan te vullen. De resultaten van de gevoerde gesprekken staan in hoofdstuk 3. Op grond van deze bronnen is bekeken op welke thema's de analyses zich het beste kunnen richten.

Vervolgens voeren we een daadwerkelijke analyse uit van mogelijk strategisch gedrag op basisscholen. Hiervoor gebruiken we vragenlijsten aan schoolleiders, leerkrachten en ouders en anderzijds grootschalige microdatabestanden. In onze empirische analyse richten we ons op de opbrengstindicatoren uit het toezichtkader, waaronder de resultaten van de eindtoets, resultaten van voortgangstoetsen, indicatoren over de ontwikkeling van leerlingen met specifieke onderwijsbehoefte. De resultaten van de empirische analyses worden beschreven in de hoofdstukken 4 en 5. Alle gebruikte vragenlijsten zijn als bijlage aan dit rapport verbonden.

Op basis van de analyses van zowel de gegevens uit de vragenlijst als de grootschalige bestanden, trekken we in hoofdstuk 6 conclusies over de effecten van strategisch gedrag bij basisscholen in Nederland en doen we aanbevelingen over de manier waarop dergelijk gedrag gemonitord en eventueel beperkt kan worden¹.

¹ Dit onderzoek is tot stand gekomen in samenwerking tussen het ITS Nijmegen en de Universiteit Maastricht (UM). Daarbij is de vraagstelling gezamenlijk geformuleerd en zijn gezamenlijk de verkennende gesprekken gevoerd. ITS heeft de enquêtering van schoolleiders en leerkrachten uitgevoerd, de UM die van ouders. De gedragseconomische analyse en de analyse van de grootschalige microbestanden en de survey resultaten zijn uitgevoerd door de UM.

2 Mogelijke vormen van strategisch gedrag

2.1 Algemene theorie over het gebruik van kwaliteitsindicatoren

De kwaliteitsindicatoren in het onderwijs hebben tot doel om vast te stellen of het onderwijs aan een school voldoende kwaliteit heeft en om scholen te stimuleren aan deze kwaliteitsnormen te voldoen. Alleen vanuit de functie om te meten of het gegeven onderwijs op een school aan de kwaliteitsnorm voldoet, ontstaan geen gedragsveranderingen. Zodra een school er echter belang bij heeft om een goed kwaliteitsoordeel van de inspectie te krijgen, zullen dergelijke indicatoren wel van invloed kunnen zijn op de gang van zaken op een school. Het kan hierbij zowel om beoogde als niet beoogde effecten gaan. Bij dat belang kan het gaan om consequenties voor de bekostiging, de beoordeling door de inspectie, maar ook het beeld dat van scholen bij ouders ontstaat. Ook kunnen indicatoren invloed hebben op de bewustwording van scholen over hun prestaties, en daarmee een stimulerende werking hebben.

Het is nuttig om drie, min of meer stapsgewijze, consequenties te onderscheiden van het gebruik van prestatie-indicatoren in het onderwijs (wettelijke normen en bekostigingsregels hebben overigens een vergelijkbaar effect). Ten eerste kunnen als reactie op de indicatoren door een school extra kosten worden gemaakt of inspanningen worden gedaan om een gunstiger beeld te creëren. Ten tweede kunnen er veranderingen in de kwaliteit van de schoolprestaties ontstaan of kunnen er kosten of baten voor anderen ontstaan, mede als gevolg van deze inspanningen. De kwaliteit en/of doelmatigheid kan daarbij verbeteren of verslechteren. Ten derde kunnen de gedragsveranderingen effect hebben op de beoordeling door de inspectie (of op de bekostiging van de school).

Zo kunnen er reacties van de school zijn op de gebruikte indicatoren die heel weinig materieel effect hebben op de onderwijskwaliteit of doelmatigheid, maar die wel van grote invloed zijn op deze indicatoren. Een voorbeeld hiervan is een leerling niet te laten deelnemen aan de eindtoets basisonderwijs als deze naar verwachting lager scoort dan 517. Als de inspectie deze leerling niet buiten beschouwing laat, dan hanteert zij de fictieve score van 517 voor deze leerling. Dit kan het gemiddelde cijfer op de eindtoets aanzienlijk verhogen, maar voor het betreffende kind heeft het weinig consequenties. Andere aanpassingen kunnen juist grote gevolgen hebben voor de kwaliteit en kosten van het onderwijs. Een voorbeeld van een kostenverhogende reactie zou zijn om leerlingen die niet goed presteren, een jaar langer over de basisschool te laten doen om zo een hoger toetsresultaat te bereiken. Om wat gunstigere indicatoren te krijgen, worden dan immers beslissingen genomen met aanzienlijke financiële consequenties. De kwaliteit van het onderwijs kan bijvoorbeeld veranderen als, vanwege de druk om goed te presteren op bepaalde toetsen, activiteiten uit het curriculum worden weggelaten die weinig invloed op die toetsen hebben. Afhankelijk van de kijk op kwaliteit kan dit een positief en een negatief effect op kwaliteit hebben. Als het gaat om activiteiten die niet als erg zinvol worden gezien, zou het een verbetering zijn. Als het gaat om vakken die wel als heel zinvol worden gezien, zou het de kwaliteit kunnen verminderen. Zo is bekend dat in de Verenigde Staten in veel scholen, onder druk van toetsen op het gebied van taal en rekenen, minder onderwijzend personeel is aangenomen voor vakken als geschiedenis en aardrijkskunde.

Veranderingen in de aanpak van een school die tot uiting komen in de gebruikte indicatoren, hoeven niet te leiden tot een ander oordeel van de inspectie. Als bijvoorbeeld door extra (gewenste of

ongewenste) inspanningen de scores op de eindtoets omhoog gaan, kan deze verhoging altijd nog onvoldoende zijn om de ondergrens die de inspectie hanteert te passeren.

Stel dat het beleid van een school heel weloverwogen is. Dan is voor die school de balans van een aanpassing van het beleid zonder dat hierbij een oordeel van de inspectie een rol speelt, de volgende (verderop gaan we in op de situatie dat er geen sprake is van een weloverwogen beleid):

Waarde van een aanpassing voor school = - Kosten (voor de school) van aanpassing beleid +/- Waarde (voor de school) van de verandering in kwaliteit

Als het beleid van de school zonder invloed van een beoordeling door de inspectie niet verandert, moet dus voor iedere beleidsverandering gelden dat deze balans negatief is. De kosten van een verandering zijn groter dan de baten er van en het huidige beleid is voor de school optimaal. Als ook de beoordeling door de inspectie een rol gaat spelen, verandert deze vergelijking in:

Waarde van een aanpassing voor school = - Kosten (voor de school) van aanpassing beleid +/- Waarde (voor de school) van de verandering in kwaliteit +/- Waarde van een goed kwaliteitsoordeel (als deze verandert door de beleidsaanpassing)

Als vanwege het inspectieoordeel een school haar beleid aanpast, moet deze tweede balans dus wel positief zijn. Deze twee vergelijkingen leveren twee interessante inzichten op.

Ten eerste, omdat het hier gaat om gedragsveranderingen die de school niet doorvoerde zonder de druk van de kwaliteitsbeoordeling, weegt kennelijk – als de school inderdaad deze afwegingen bewust maakt – de waarde voor de school van de verandering in kwaliteit op zichzelf niet op tegen de kosten van de aanpassing van het beleid. Met andere woorden, de balans van de eerste vergelijking zal voor de school negatief zijn, anders had zij het ook zonder de invloed van het kwaliteitsoordeel doorgevoerd. Ten tweede, de totale balans kan alleen positief zijn als het kwaliteitsoordeel van de inspectie door het veranderde beleid verandert. Uiteraard zal toeval een rol spelen bij dit soort processen, dus in ieder geval betekent dit dat een kwaliteitsbeoordeling alleen tot gedragsveranderingen bij de school zal leiden, als de school inschat dat de kans op een verbetering in het oordeel groot genoeg is. Een consequentie van dit tweede inzicht is dat scholen die reeds ruim boven een ondergrens scoren of scholen die daar ver onder zitten en met een aanpassing beter kunnen gaan scoren maar niet boven deze grens zullen komen, niet geprikkeld worden om het beleid aan te passen.

Voor de inspectie is een andere balans van toepassing:

Waarde van een aanpassing voor de inspectie = - Kosten (voor de inspectie) van aanpassing beleid +/- Waarde (voor de inspectie) van de verandering in kwaliteit

Als we er van uit gaan dat voor de inspectie zelf het kwaliteitsoordeel geen kosten met zich meebrengt, tellen voor haar alleen de kosten van het beleid en waarde van de kwaliteitsverandering, die daar het gevolg van is. Een vergelijking van de balans voor de school en die van de inspectie levert opnieuw twee belangrijke inzichten op.

Ten eerste, als de gedragsveranderingen die veroorzaakt worden door de kwaliteitsbeoordeling gunstig uitpakken in de ogen van de inspectie, moet het dus zo zijn dat de gerealiseerde

kwaliteitsverandering in de ogen van de inspectie gunstiger is dan de kwaliteitsverandering in de ogen van de school. Met andere woorden, kwaliteitsoordelen die van invloed zijn op het gedrag van scholen kunnen alleen een gunstig gedragseffect hebben, als het beleid dat optimaal is in de ogen van de school, afwijkt van het beleid dat volgens de inspectie optimaal is. Ten tweede, en eigenlijk de keerzijde hiervan, als de school en de inspectie kwaliteitsverbeteringen en de kosten die hiervoor gemaakt moeten worden, op dezelfde wijze waarderen, kan de gedragsverandering als gevolg van het gebruik van kwaliteitsindicatoren alleen maar leiden tot een verslechtering van de kwaliteit van het onderwijs. In dit geval streeft de school feitelijk hetzelfde na als de inspectie maar heeft de beoordeling een frustrerende invloed op dit gewenste beleid. Als het onderwijs bewust haar afwegingen maakt, heeft een beoordelingskader dus alleen zin als scholen zelf andere afwegingen zouden maken bij het bevorderen van goed onderwijs dan de inspectie zou doen.

Als scholen reageren op de indicatoren die gehanteerd worden door de inspectie door meer inspanningen te leveren gericht op de verbetering van de prestaties van de leerlingen, is er sprake van gewenst strategisch gedrag in lijn met het doel van de beoordeling. Een school kan zich echter ook richten op verbetering van de beeldvorming, met in het uiterste geval negatieve consequenties voor de onderwijskwaliteit en de leerprestaties. In de praktijk zullen zich vaak tussenvarianten voordoen, waarbij de extra inspanningen weliswaar de onderwijsopbrengsten verbeteren, maar waarbij met dezelfde inspanningen ook een beter resultaat had kunnen worden geboekt als deze exclusief zouden zijn gericht op de verbetering van de kwaliteit van het onderwijs en niet op verbetering van de beeldvorming. De indicatoren hebben dan een gunstig stimulerend effect gehad, maar met daarbij een efficiencyverlies. In zo'n geval zouden indicatoren die in hoge mate scholen aanzetten tot kwaliteitsverbetering, en waarbij minder efficiencyverlies optreedt, beter werken, maar kan het gebruik van deze indicatoren beter zijn dan het niet gebruiken van deze indicatoren.

De invloed van kwaliteitsindicatoren in het onderwijs is vergelijkbaar met de invloed van incentives voor managers in het bedrijfsleven. In de zogenaamde *management control* literatuur wordt aangegeven dat goede kwaliteitsindicatoren moet voldoen aan de volgende eisen²: Congruentie of de indicatoren in lijn zijn met de doelen die moeten worden beoogd, kennis van de beoogde resultaten, de mogelijkheid om de beoogde resultaten te beïnvloeden, de mogelijkheid om de beoogde resultaten vast te stellen, nauwkeurigheid, objectiviteit, tijdigheid en begrijpelijkheid.

De ongewenste effecten van het gebruik van de indicatoren kunnen worden onderscheiden in plaatsvervangend gedrag (*behavioral displacement*), tactisch handelen (*gamesmanship*) en negatieve attitudes (*negative attitudes*)³. Van plaatsvervangend gedrag is sprake als de gehanteerde indicatoren niet congruent zijn met de werkelijke doelen van de inspectie. De scholen gaan zich dan inspannen om resultaten te behalen die eigenlijk niet beoogd zijn. Van tactisch handelen is sprake als een school zich in gaat spannen om indicatoren gunstig te beïnvloeden, zonder dat dit van invloed is op de feitelijke schoolprestaties. Van negatieve attitudes is sprake als de indicatoren leiden tot frustratie, spanningen of conflict.

Als de indicatoren, zoals gemeten, geen valide weergave van de daadwerkelijk onderliggende onderwijskwaliteit zijn, kan strategisch gedrag er ook toe leiden dat de inspectie verkeerde conclusies

² Merchant, K.A. en Van der Stede, W. (2007), *Management Control: Measurement, evaluation and incentives*. Pearson Education Limited.

³ K.A. Merchant (1998), *Modern Management Control Systems*. Prentice Hall.

trekt. Dit kan uiteenlopende oneigenlijke effecten met zich meebrengen, zoals ouders die een school kiezen op basis van onjuiste informatie, resulterend in een grotere toestroom van leerlingen en meer middelen. Ook zal het beeld van het onderwijs op stelselniveau vertroebeld raken door de wijze waarop indicatoren het gedrag van scholen beïnvloeden.

Economische theorieën bieden een goede basis om te voorspellen welk strategisch gedrag bepaalde indicatoren uit zullen lokken. Als een school als enige doel heeft gunstig te scoren op de indicatoren, kan geanalyseerd worden op welke manier dit met zo min mogelijk inspanningen kan worden bereikt. In de praktijk worden mensen uiteraard ook gedreven door intrinsieke waarden en spelen andere (soms impliciete) mechanismen een rol die het gedrag beïnvloeden. In de vergelijkingen hierboven betekent dit dat de school niet puur gedreven zal worden door de consequenties van haar handelen voor het kwaliteitsoordeel. De geleverde kwaliteit als zodanig is ook een belangrijke drijfveer voor de betrokkenen. Het gevolg hiervan is in veel gevallen dat het strategisch gedrag optreedt in de situaties die de extreme theorie (die waarbij de geleverde kwaliteit als zodanig geen drijfveer is) voorspelt, maar dat de mate waarin dit optreedt kleiner is.

In het algemeen zijn er vier vormen van gedragsverandering van scholen in reactie op kwaliteitsindicatoren mogelijk:

1. De indicatoren beperken scholen in de mogelijkheid om goed beleid te voeren.
2. De indicatoren stimuleren scholen om kwaliteit te leveren.
3. De indicatoren stimuleren scholen om meer aandacht te besteden aan de gemeten indicatoren en minder aan de niet-gemeten indicatoren.
4. De indicatoren stimuleren scholen om meer aandacht te besteden aan de gemeten indicatoren zonder dat dit de kwaliteit ten goede komt.

Deze vier vormen kunnen overigens ook in combinatie met elkaar voorkomen. Hieronder bespreken we de vier vormen echter afzonderlijk om de mechanismen duidelijk te maken.

De indicatoren beperken scholen in de mogelijkheid om goed beleid te voeren

Stel dat een school exact dezelfde intrinsieke doelen heeft als de inspectie zou willen. In dat geval kunnen indicatoren alleen belemmerend werken (of geen effect hebben). Dat probleem zal zich met name voordoen als de school over informatie beschikt die de inspectie niet heeft en dus niet in haar beoordeling kan betrekken. Een voorbeeld is dat een school van iedere leerling bij binnenkomst zijn gemiddelde potentie kan beoordelen op basis van opgevraagde informatie over de leerling, maar dat de inspectie niet over deze informatie beschikt en zich hierbij moet richten op ruwe schattingen hiervoor, zoals het gemiddelde opleidingsniveau van de ouders. Deze beperkende invloed van de kwaliteitsindicatoren leidt tot frustratie bij de schoolleiders en leerkrachten. Als werkelijk alle scholen optimaal zouden functioneren en overschrijdingen van de normen alleen het gevolg van toeval zouden zijn, zouden de normen alleen kwaliteitsverlagend en frustratieverhogend werken. Als er echter ook een gunstige stimulerende werking van het gebruik van indicatoren uitgaat, ontstaat er een lastige trade-off. De positieve invloed zal dan afgewogen moeten worden tegen deze negatieve effecten. Gekeken kan worden naar middelen om te voorkomen dat scholen ongewenst worden ingeperkt in hun handelingsruimte. Te denken valt hierbij aan gesprekken tussen de inspectie en scholen naar aanleiding van beoordelingen of systemen waarbij scholen onderling peer review organiseren. De trade-off kan echter nooit helemaal worden opgeheven.

De indicatoren stimuleren een school kwaliteit te leveren

Tegenover deze inperkende werking van indicatoren staat een stimulerende werking. Als een school zonder toezicht minder prestaties levert dan de inspectie redelijk acht, kunnen indicatoren een stimulerende werking hebben. Dat kan via twee mechanismes lopen:

In de eerste plaats kan het zijn dat de school geen goed beeld heeft van haar opbrengsten. De informatie van de indicator maakt dan helder dat de prestaties onder de maat zijn, wat een motivatie voor de medewerkers kan zijn om voor verandering te zorgen. In de tweede plaats kan de indicator ook een stimulans vormen, als er consequenties voor de school zijn verbonden aan de gemeten indicatoren.

Deze consequenties kunnen diverse vormen aannemen. De betrokken medewerkers kunnen een negatief oordeel van de inspectie als vervelend ervaren, het kan zijn dat het invloed heeft op het keuzegedrag van ouders en er kunnen sancties verbonden zijn aan het oordeel.

Een beoordeling die gebaseerd is op een indicator die een goed beeld geeft van de kwaliteit van een school vergroot daarmee de waarde voor een school om goede prestaties te leveren. De feitelijke inspanningen van een school zullen altijd afhangen van een afweging van de kosten (geld, hoeveelheid werk) die gemaakt moeten worden om bepaalde prestaties te leveren en de waardering van die opbrengst. Indicatoren kunnen de waardering van de opbrengst verhogen, waardoor deze afweging meer ten gunste van de prestaties uit zal vallen.

Een groot probleem bij deze stimulerende werking van indicatoren is overigens dat het voor de inspectie moeilijk zal zijn om per school vast te stellen wat een redelijke prestatie is. Er zullen immers grote verschillen zijn tussen scholen in de mogelijkheden die zij hebben, bijvoorbeeld gezien de instroom van leerlingen. Geprobeerd kan worden om voor scholen die in verschillende omstandigheden verkeren verschillende normen te hanteren. Als er verschillen overblijven waar geen rekening mee wordt gehouden, of kan worden gehouden, kan een situatie die voor verschillende scholen verschillend uitpakt. Voor de ene school kan het heel makkelijk zijn om aan de norm te voldoen. Voor een andere school kan een extra inspanning nodig zijn om de norm te halen. En voor weer een andere school kan de norm praktisch onhaalbaar zijn. De consequentie is dat met name voor de tweede groep de norm een stimulans zal vormen om prestaties te verhogen.

De indicatoren stimuleren scholen om meer aandacht te besteden aan de gemeten indicatoren en minder aan de niet-gemeten indicatoren

De hierboven besproken vormen van invloed van indicatoren op het beleid van scholen zijn impliciet gebaseerd op de gedachte dat kwaliteit eindimensionaal is. Scholen kunnen zich meer of minder inspannen om hier iets aan te doen en de indicatoren hebben daar een invloed op. De output van het onderwijs is echter meerdimensionaal. Een gevolg daarvan is dat de stimulans om in één of enkele dimensies beter te scoren, niet per se tot gevolg hoeven te hebben dat de school meer investeert of meer inspanningen doet, maar dat het er ook toe kan leiden dat de inspanningen op andere fronten worden verminderd.

De interactie tussen inspanningen voor verschillende aspecten van het onderwijs is overigens zeer complex. Allereerst kan het zo zijn, dat maatregelen die een school neemt om beter te presteren op de terreinen die wel gemeten worden met een indicator, ook gunstig uitpakken voor de terreinen die niet worden gemeten. Gedacht kan worden aan algemene kwaliteitsbevorderende maatregelen, zoals meer

structuur in het lesprogramma. Ook kunnen verbeteringen in de leerprestaties op het terrein van het ene vak (bijvoorbeeld taal) een positieve uitwerking hebben op het leren in andere vakgebieden. Ten slotte kunnen verschillende leerstoffen elkaar aanvullen, of verschillende onderdelen van de leerstof elkaars substituut zijn. In het geval van substituten, maakt het voor het succes in het vervolgonderwijs niet veel uit of een leerling juist veel van de ene vaardigheid of veel van de andere vaardigheid beheerst. In dat geval hoeft een focus van de indicatoren op het ene terrein geen negatieve gevolgen te hebben. Leerstoffen kunnen echter ook complementair zijn. In dat geval hebben de verworven vaardigheden vooral waarde in het vervolgonderwijs als ze beide op een vergelijkbaar niveau ontwikkeld zijn. Als met de indicatoren één van de twee vakgebieden wordt gestimuleerd, betekent dit dat de meerwaarde van het andere vakgebied zal gaan toenemen. Als er naast de werking van de incentives ook nog voldoende andere krachten zijn die sturen in hoeverre leerlingen leren wat ze zouden moeten leren, zullen die krachten sterker de nadruk gaan leggen op de vakken die juist niet met indicatoren worden gestimuleerd.

De indicatoren stimuleren scholen om meer aandacht te besteden aan de gemeten indicatoren zonder dat dit de kwaliteit ten goede komt

Naast de mogelijkheid om verschuivingen aan te brengen in de aandacht die scholen geven aan bepaalde vakgebieden, zijn er in de praktijk talloze manieren waarop het onderwijs indicatoren kan verbeteren, zonder dat dit gepaard gaat met een daadwerkelijke kwaliteitsverbetering van het onderwijs. Hierbij kan gedacht worden aan het verbeteren van de omstandigheden in het toetslokaal, het bevorderen dat de leerlingen fit en geconcentreerd op de toets verschijnen, leerlingen niet laten deelnemen aan de toets, het geven van instructies tijdens de toets of het aanpassen van antwoorden na de toets. Hiermee gaan de toetsresultaten omhoog, zónder dat dit samen hoeft te gaan met een toename van de vaardigheid die de toets meet. Het kan hier gaan om gedragingen die toelaatbaar zijn, maar het kan ook gaan om fraude.

Deze analyse van de effecten van kwaliteitsindicatoren op het beleid van scholen is gebaseerd op de veronderstelling dat scholen een weloverwogen beleid voeren en keuzes maken die hun doelen zo veel mogelijk realiseren. In de praktijk kunnen beoordelingen van de inspectie natuurlijk ook een signaalfunctie hebben. Wellicht heeft een school zelf niet in de gaten dat haar kwaliteit onder de maat is en schudt een oordeel van de inspectie de organisatie wakker. Van strategisch gedrag is dan geen sprake. Eerder van passief gedrag. De consequenties van toezicht op scholen is dan dus dat scholen bewust worden gemaakt van de afwegingen die ze maken. Dit bewustwordingsproces is met name van toepassing op de aspecten die de inspectie in haar beoordeling betreft. Naast de positieve en negatieve effecten van strategisch gedrag kan de bewustwording worden gezien als een effect van het toezicht.

2.2 Literatuur over het gebruik van kwaliteitsindicatoren

Studies tonen aan dat het gebruik van kwaliteitsindicatoren positieve effecten kan hebben op de effectiviteit van scholen⁴. Het gebruik van de indicatoren speelt een cruciale rol in het proces van

⁴ Hanushek, E.A. & M.E. Raymond (2004). Does school accountability lead to improved student performance? NBER-working paper no. 10591, Massachusetts. Wij beperken ons hier tot de literatuur na 2000. Een uitgebreid literatuuroverzicht

opbrengstgericht werken en accountability en maakt benchmarking met andere scholen mogelijk⁵. Zo kan feedback op de uitkomsten van de kwaliteitsindicatoren, bijvoorbeeld door middel van gesprekken tussen de inspecteur en de schooldirectie over de zwakke onderdelen van het onderwijs, of via het laten opstellen van verbeterplannen, tot positieve actie en onderwijsverbeteringen leiden⁶. De onderwijsverbetering wordt met name gevonden als het gaat om relatief kleine aanpassingen in het beleid (zoals wijzigingen in het toetsbeleid)⁷. Andere studies laten echter zien dat feedback op de kwaliteitsindicatoren soms moeilijk te interpreteren is voor schoolleiders en leerkrachten en er daarom niets mee wordt gedaan⁸. Enkele studies laten in het jaar van het inspectiebezoek een negatief effect op de onderwijsuitkomsten zien, waarbij de rol van de stress en inspanning met betrekking tot de voorbereiding van het bezoek als belangrijkste verklaring wordt gezien⁹. Een aantal studies laat de ervaringen zien die schooldirecteuren (met name in Nederland en Engeland) hebben met de inspectiebezoeken. De meerderheid van de schooldirecteuren is tevreden met de inspectiebezoeken en meent dat deze een positief effect hebben op de kwaliteit van het onderwijs. Deze positieve mening hangt overigens sterk samen met het uiteindelijke oordeel van de inspectie¹⁰.

Veel onderzoek naar de effecten van kwaliteitsindicatoren richt zich op het gebruik van ranglijsten waarin de resultaten van scholen worden weergegeven en openbaar worden gemaakt voor publiek. Het opstellen van zulke ranglijsten is niet zonder methodologische problemen en de effecten ervan op schoolverbetering zijn niet duidelijk¹¹. Ook de effecten op het schoolkeuzegedrag van ouders zijn niet eenduidig, waarbij enkele auteurs wijzen op het feit dat ouders deze informatie niet kunnen vinden of niet weten hoe deze te interpreteren¹². Recente experimenten laten zien dat als ouders beter

dat ook ingaat op literatuur van voor 2000 is te vinden in de Wolf, I.F. & F.J.G. Janssens (2007). Effects and side effects of inspections in accountability and education: an overview of empirical studies. Working paper Amsterdam University / Dutch Inspectorate of Education, Amsterdam.

⁵ Fryer, K., J. Antony & S. Ogden (2009). Performance management in the public sector. *International Journal of Public Sector Management*, 22(6): 478-498.

Fryer, R. (2012). Injecting successful charter school strategies into traditional public schools: early results from an experiment in Houston, NBER working paper no. 17494, Massachusetts.

⁶ Chapman, C. (2001). Changing classrooms through inspection. *School Leadership and Management*, 21(1): 59-73.

Ehren, M.C.M. & A. Visscher (2008). The relationships between school inspections, school characteristics and school improvement. *British Journal of Educational Studies*, 56(2): 205-227.

⁷ Ehren, M.C.M. & A. Visscher (2008). The relationships between school inspections, school characteristics and school improvement. *British Journal of Educational Studies*, 56(2): 205-227.

⁸ Hellrung, K. & J. Hartig (2013). Understanding and using feedback: A review of empirical studies concerning feedback from external evaluations to teachers.

Verhaeghe, J. P. (2010). Schoolfeedback als input voor interne kwaliteitszorg. In: *Handboek Reflectief Vermogen*.

Visscher, A. J. & R. Coe (2003). School performance feedback systems : conceptualisation, analysis, and reflection. *School effectiveness and school improvement*, 14 (3): 321-349.

⁹ Rosenthal, L. (2004). Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. *Economics of Education Review*, 23: 143-151.

Shaw, I., D.P. Newton, M. Aitkin, & R. Darnell (2003). Do Ofsted inspections of secondary education make a difference to GCSE results? *British Educational Research Journal*, 29(1): 63-75.

¹⁰ Emmelot, Y., S. Karsten, G. Ledoux, & A. Vermeulen (2004). *Ervaringen met het vernieuwde onderwijs-toezicht*. Amsterdam: SCO/Kohnstamm Instituut.

Mathews, P. & P. Sammons (2004). *Improvement through Inspection. An evaluation of the impact of Ofsted's work*. London: Ofsted.

¹¹ Goldstein, H. (2001). League tables and schooling. *Science in Parliament*, 58(2): 4-5.

Kane, T.J. & D.O. Staiger (2001) *Improving school accountability measures*. NBER working paper no. 8156. Massachusetts.

¹² Janssens, F.J.G. & A.J. Visscher (2004). Naar een kwaliteitskaart voor het primair onderwijs. *Pedagogische Studiën*, 81: 371-383.

Visscher, A.J. (2004). De effecten van schoolprestatie-publicaties in andere landen. *Pedagogische Studiën*, 81: 384-396.

geïnformeerde keuze maken, dit een positief effect heeft op de schoolprestaties van de leerlingen¹³. Nederlandse schooldirecteuren geven aan dat de publicatie van kwaliteitsindicatoren van scholen onderwijsverbetering stimuleert¹⁴.

Er worden echter ook risico's en negatieve effecten van het gebruik van kwaliteitsindicatoren in het onderwijs gerapporteerd in de literatuur¹⁵. Zo bestaat het risico dat de kwaliteit van de school en het onderwijs wordt gereduceerd tot zaken die zich laten meten¹⁶. Scholen leveren een brede bijdrage aan de leerontwikkeling van leerlingen, maar niet alle vaardigheden laten zich even makkelijk meten. Dit kan ook betekenen dat er *teaching to the test* plaatsvindt¹⁷, of dat de leertijd wordt gereduceerd doordat meer tijd aan toetsen wordt besteedt en er een focus ligt op de verantwoording in plaats van 'leren'¹⁸. Tot slot kan het gebruik van kwaliteitsindicatoren leiden tot gedrag van betrokken scholen/leerkrachten, waardoor de scores op de indicatoren gunstiger zijn dan de onderliggende kwaliteit rechtvaardigt¹⁹.

2.3 Theorie over strategisch gedrag

Strategisch gedrag is in de meeste gevallen niet goed te scheiden van regulier handelen. Vaak zijn de aanpassingen in het gedrag graduëel. Zo ligt het voor de hand om leerlingen als voorbereiding op de eindtoets toetsen uit eerdere jaren te laten maken, om hen een beeld te geven van wat ze kunnen verwachten. Vanwege de belangen die er op het spel staan, zou een school dit echter meer en meer kunnen gaan doen, waardoor geleidelijk aan normaal gedrag verandert in strategisch gedrag. Een ander voorbeeld zou de beslissing kunnen zijn, om een leerling toe te laten tot de eigen school of door te verwijzen naar speciaal onderwijs. Het ligt niet voor de hand dat uit strategische overwegingen leerlingen worden doorverwezen die eigenlijk gewoon op de school waar ze worden aangemeld thuishoren, maar in de praktijk zullen er veel grensgevallen zijn waarbij een inschatting gemaakt moet worden wat de beste keuze is. In zo'n situatie kan bewust of onbewust meespelen welke consequenties dit voor het oordeel van de inspectie heeft.

Het ligt dus voor de hand dat scholen voor een groot deel handelen vanuit een eigen visie over wat goed onderwijs is en dat deze manier van handelen deels beïnvloed wordt door de kwaliteitsindicatoren van de inspectie. Er kan echter ook een andere invloed van indicatoren op het beleid van scholen zijn. Het is in het basisonderwijs niet vanzelfsprekend wat goed onderwijs is. De kwaliteit van het onderwijs hangt feitelijk af van de betekenis van dit onderwijs voor de latere loopbaan van leerlingen en dit onttrekt zich grotendeels aan de waarneming van de schoolleiders en leerkrachten.

¹³ Hastings, J.S., T.J. Kane & D.O. Staiger, D. (2010). Heterogeneous Preferences and the Efficacy of Public School Choice. NBER Working Paper no. 12145, Massachusetts.

Hastings, J. S., R. van Weelden & J. Weinstein, J. (2007). Preferences, Information and Parental Choice Behavior in Public School Choice. NBER Working Paper no 12995, Massachusetts.

¹⁴ Emmelot, Y., S. Karsten, G. Ledoux, & A. Vermeulen (2004). Ervaringen met het vernieuwde onderwijstoezicht. Amsterdam: SCO/Kohnstamm Instituut.

¹⁵ Wolf, I. de en M. Swanborn (2008). Betrouwbaarheid en validiteit van opbrengstmaten in het primair onderwijs. Paper voor de Onderwijsresearchdagen in Eindhoven.

¹⁶ Ledoux, G., H. Blok, & M. Boogaard (2009). Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs. Amsterdam: SCO-Kohnstamm Instituut.

¹⁷ Lazear, E.P. (2006). Speeding, terrorism and teaching to the test. Quarterly Journal of Economics, 121(3): 1029-1061.

¹⁸ Ledoux, G., H. Blok, & M. Boogaard (2009). Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs. Amsterdam: SCO-Kohnstamm Instituut.

¹⁹ Koretz, D. M. (2002). Limitations in the use of achievement tests as measures of educators' productivity. Journal of human resources, 752-777.

Het beeld wat goed onderwijs is, wordt daarom waarschijnlijk ook gevormd door gewoonte, verwachtingen of persoonlijke opvattingen. Indicatoren zullen hier ook invloed op hebben. Uit de gedragseconomie is bekend dat zaken die een detail kunnen lijken, grote invloed hebben op het voorkomen van strategisch gedrag omdat ze de interpretatie van de setting en daarmee de intrinsieke motivatie in hoge mate beïnvloeden. Als gedurende lange tijd een bepaalde indicator gebruikt wordt om de kwaliteit van het onderwijs te meten, bestaat de kans dat voor betrokkenen deze indicator geleidelijk synoniem met kwaliteit gaat worden.

Een klassiek voorbeeld hiervan is de studie van de kinderopvang in Israël²⁰. Omdat veel ouders hun kind vaak te laat kwamen ophalen, voerde de zaal een boete in voor te laat komen. Dit had een enorme toename in het laatkomen tot gevolg. De ouders hadden de indruk dat zij met de boete een vergoeding betaalden voor de gevolgen van hun gedrag en werden daardoor niet meer geremd door schuldgevoelens tegenover de leidsters die langer moest wachten. Toen de maatregel werd teruggedraaid keerde het oude gedrag echter niet terug. De tijdelijke invoering van een boete had definitief de mindset van de mensen veranderd. Experimenten laten ook zien dat het gedrag van anderen van grote invloed is op het gedrag van het individu²¹. In een experiment waarin mensen kunnen kiezen tussen zelfzuchtig of sociaal gedrag, kiest bijna altijd een deel van de mensen voor het zelfzuchtig gedrag en een deel van de mensen voor het sociaal gedrag. Het experiment laat zien dat als mensen in een omgeving zijn waarin veel anderen zelfzuchtig zijn, zij ook sneller hun sociale gedrag op zullen geven: Als anderen alleen maar aan zichzelf denken, dan doe ik dat ook. Hierdoor kan er, door een kleine prikkel, een snelle omslag plaatsvinden in het typisch gedrag van een groep doordat aanvankelijk enkelen reageerden maar daardoor uiteindelijk anderen ook hun gedrag gaan aanpassen. Uit een andere studie naar bloeddonthaties blijkt dat extrinsieke motivatie de intrinsieke motivatie kan verdringen als de manier waarop waardering wordt gegeven verandert²².

Een recent experiment illustreert hoe mensen zich gedragen als ze niet worden geobserveerd²³. Dit experiment is een ijkpunt in de gedragseconomie geworden voor methodes om strategisch gedrag op te sporen. In dit experiment mogen participanten met een dobbelsteen gooien met de cijfers 0 tot en met 5. De hoogte van hun worp bepaalt het bedrag dat zij krijgen, bijvoorbeeld bij 0 krijgen ze 0 euro, bij 1 krijgen ze 1 euro, ... en bij een worp van 5 krijgen ze 5 euro. De condities worden zo opgezet dat de participant zijn dobbelsteen kan werpen, zonder dat hij wordt geobserveerd en waarbij hij ook zeker kan weten dat hij niet geobserveerd kan worden. Hij wordt gevraagd om de hoogte van zijn worp te noteren en na afloop is deze zelf opgegeven score bepalend voor de betaling.

Figuur 2.1 geeft de verdeling van de opgegeven worpen bij het experiment. Een aantal zaken is interessant. Allereerst maakt de grafiek duidelijk dat een deel van de participanten niet hun werkelijk worp doorgeeft. Bij een juiste opgave zouden alle mogelijke worpen ongeveer even vaak moeten voorkomen, maar de hoge cijfers worden veel vaker doorgegeven dan de lage cijfers. In de tweede plaats laat de grafiek zien, dat sommige deelnemers wel een juiste opgave van zaken doen, ook als dat niet in hun belang is. Er is namelijk een redelijke groep participanten die aangeeft 0 te hebben

²⁰ Gneezy, U. & A. Rustichini (2000). Fine Is a Price, *American Journal of Legal Studies*, 29(1).

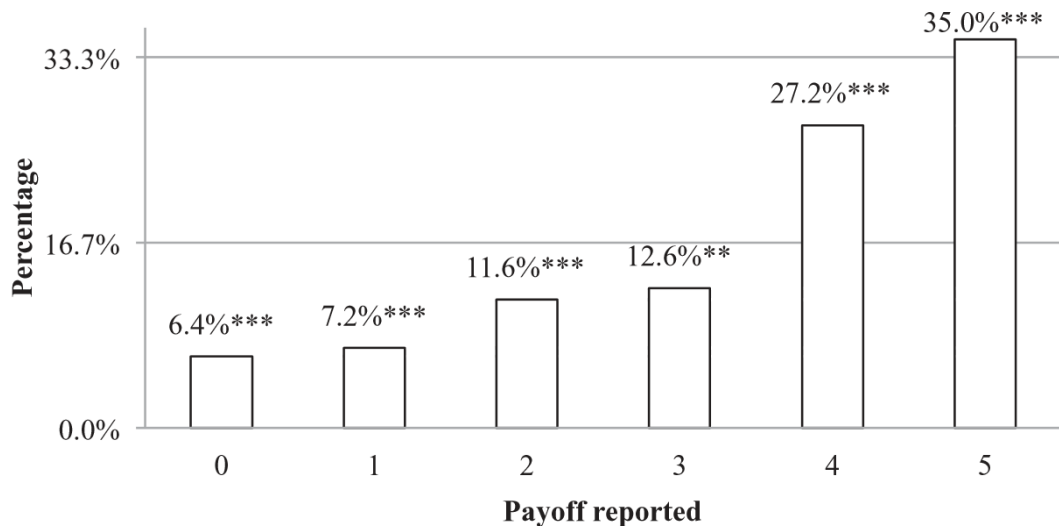
²¹ Fehr, E. & U. Fischbacher (2003). The nature of human altruism. *Nature* 425.6960 (2003): 785-791.

²² Mellström, C. & M. Johannesson (2008). Crowding out in blood donation: was Titmuss right? *Journal of the European Economic Association*, 6(4): 845-863.

²³ Fischbacher, U. & F. Föllmi-Heusi (2013). Lies in Disguise—an Experimental Study on Cheating. *Journal of the European Economic Association*, 11(3): 525-547.

geworpen. Redelijkerwijs kan worden aangenomen dat zij hun worp correct hebben doorgegeven. In de derde plaats is het opmerkelijk, dat degenen die geen waarheidsgetrouwe opgave doen, niet per se de maximale winst proberen binnen te halen. Er zijn immers ook meer vieren dan men statistisch gezien zou verwachten en dus geeft een deel van de participanten een te hoog getal op, maar beperken ze dat tot een lager getal dan vijf.

Figuur 2.1 Opgegeven betaling in het experiment van Fischbacher en Föllmi-Heusi (2013)



Het experiment laat zien dat er sprake is van strategisch gedrag, waarbij sommige participanten zich strategisch gedragen en anderen niet, en waarbij sommige participanten die zich strategisch gedragen dit ook op de manier doen die de grootste winst oplevert, maar anderen dat slechts gedeeltelijk doen. De mate waarin mensen zich strategisch gedragen zal afhangen van de omvang van de stimulans en van de specifieke situatie.

Tegelijkertijd geeft het experiment ook de kern van de oplossing van het dilemma hoe strategisch gedrag vastgesteld kan worden. Los van “betrappen” komt zo’n aanpak er altijd op neer dat een situatie wordt gecreëerd waarin op individueel niveau geen consequenties kunnen worden verbonden aan informatie van een betrokkene, maar op geaggregeerd niveau toch duidelijk wordt wat er aan de hand is. Het probleem van betrappen is vaak, dat als eenmaal duidelijk is op welke wijze scholen betrap worden op ongewenst strategisch gedrag, men het beleid zodanig aan gaat passen dat het in het vervolg niet meer eenvoudig gedetecteerd zal worden. Het aanpakken van ongewenst strategisch gedrag leidt daarmee tot nieuwe vormen van strategisch gedrag. Deze kan alleen tot staan worden gebracht als alle mogelijkheden om indicatoren te verhogen zonder daarbij het beoogde gedrag te vertonen zijn uitgeput.

2.4 Strategisch gedrag in de literatuur

De literatuur laat zien dat scholen strategisch reageren op de introductie van stimulansen die gerelateerd zijn aan leerlingprestaties. De gedragseffecten van het gebruik van de indicatoren van de inspectie komen in hoge mate overeen met de gedragseffecten van prestatiebekostiging. Uit de literatuur komen veel voorbeelden naar voren van dergelijk strategisch gedrag, zowel in het onderwijs

als daarbuiten²⁴. Zo worden gedragseffecten gevonden die positief zijn voor het leerproces van leerlingen en de leerprestaties verbeteren. Het gaat daarbij om scholen die hun onderwijsaanbod vergroten, effectievere instructie geven, of de instructietijd in de vakken waarin leerlingen zwak presteren²⁵. Andere studies laten zien dat sommige scholen in de Verenigde Staten meer leerlingen aanmerken als zwakke of zorgleerlingen zodat ze uitgesloten kunnen worden van de toetsen²⁶. De percentages leerlingen die waren uitgesloten van de test verdubbelden in sommige gevallen van 7 naar 14 procent. Op andere Amerikaanse scholen bleek het verschil tussen schorsing van zwakke leerlingen en goede leerlingen in toetsperiodes voor schoolbeoordeling (high-stakes toetsen in de VS) significant groter dan gedurende de rest van het jaar of gedurende andere toetsperiodes²⁷.

Een studie naar de effecten van veranderingen in het toezicht naar aanleiding van de “*No Child Left Behind*”-act onder publieke scholen in Chicago laat zien dat schooldirecteuren en leraren strategisch reageren op de veranderende stimulans, op verschillende manieren²⁸: (1) er werden meer plekken gecreëerd voor speciaal onderwijs op de scholen; (2) er bleven meer leerlingen zitten; (3) er werd minder aandacht besteedt aan vakken die niet terug kwamen op de toetsen die onderdeel uitmaakten van het toezicht (waaronder science en sociale vakken). Het verschuiven van de aandacht van vakken die niet gemeten worden op de toets naar vakken die wel gemeten worden op de toets komt overigens ook in andere studies naar voren²⁹. Andere studies laten zien dat scholen gaan oefenen met de test, of met toetsen afnemen in het algemeen, het geen de testcores wellicht kunstmatig verhoogt³⁰.

Andere voorbeelden van ongewenst strategisch gedrag, in het bijzonder fraude, zijn het voorzeggen van de juiste antwoorden tijdens de toets, het aanpassen van antwoorden door leerkrachten achteraf, of de opgaven van de toets een dag voor afname van de toets met de leerlingen doornemen³¹.

²⁴ Koretz, D. M., & S.I. Barron (1998). The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS).

Koretz, D. M. (2002). Limitations in the use of achievement tests as measures of educators' productivity. *Journal of Human Resources*, 752-777.

²⁵ Stecher, B.M. (2002). Consequences of large-scale, high-stakes testing on school and classroom practices. Tests and their use in test-based accountability systems. In Hamilton, L.S., B.M. Stecher & S.P. Klein (Eds.). *Making sense of Test-based Accountability in Education*. Santa Monica: Rand cooperation.

Koretz, D.M., D.F. McCaffrey, & Hamilton, L.S. (2001). Towards a Framework for Validating Gains under High-Stakes Conditions. CRESST/Harvard Graduate School of Education: CSE Technical Report 551.

²⁶ Cullen, J.B. & R. Reback (2006). Tinkering Toward Accolades: School Gaming under a Performance Accountability System, in T. J. Gronberg & D. W. Jansen (ed.) *Improving School Accountability (Advances in Applied Microeconomics volume 14)* Emerald Group Publishing Limited, pp.1 – 34.

Figlio, D.N. & L.S. Getzler (2006). Accountability, ability and disability: Gaming the system? in T. J. Gronberg & D. W. Jansen (ed.) *Improving School Accountability (Advances in Applied Microeconomics volume 14)* Emerald Group Publishing Limited, pp. 35-49.

²⁷ Figlio, D. N. (2006). Testing, Crime And Punishment. *Journal of Public Economics*, 90(4-5): 837-851.

²⁸ Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of public Economics*, 89(5): 761-796.

²⁹ Koretz, D. M., & S.I. Barron (1998). The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS).

³⁰ Holcombe, R., J. Jennings & D. Koretz (2013). The roots of score inflation: An examination of opportunities in two states' tests. In G. Sunderman (Ed.), *Charting reform, achieving equity in a diverse nation*, 163-189. Greenwich, CT: Information Age Publishing.

Klein, S.P., L.S. Hamilton, D.F. McCaffrey & B.M. Stecher (2000). What do test scores in Texas tell us? Issue Paper, Rand Education, California.

Stecher, B. & S.I. Barron (1999). *Quadrennial Milepost Accountability Testing In Kentucky*. Center for the Study of Evaluation, University of California, Los Angeles.

Koretz, D. M., & S.I. Barron (1998). The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS).

³¹ Amrein-Beardsley, A., D.C. Berliner, S. Rideau (2010). Cheating in the first, second and third degree: Educators' responses to high-stakes testing. *Education Policy Analysis Archives*, 18(14).

Om fraude op te sporen gebruikten enkele onderzoekers data van publieke scholen in Chicago en ontwikkelden een algoritme³². Dit algoritme ging na of er ongewone patronen zichtbaar werden in de antwoorden op de toetsvragen. Figuur 2.2 laat zien dat er patronen zichtbaar zijn die het zeer aannemelijk maken dat er fraude heeft plaatsgevonden. Hele blokken test scores (dus antwoordreeksen van verschillende studenten) zijn identiek en correct. De studenten in deze klas hebben ook opvallend hogere toetsscores dan in voorgaande jaren. Een dergelijke fraude werd in ongeveer vier tot vijf procent van de klassen gevonden. Het werkelijke percentage fraude kan hoger zijn omdat het algoritme geen andere vormen van fraude oppikt, zoals studenten meer tijd geven voor de toets of het weglaten van zwakke leerlingen uit de toets.

Figuur 2.2 Fraude op de toets, ontdekt door Jacob en Levitt (2003)

Student Answer Strings (each row represents one student's answers)	Student Test Scores		
	Year t-1	Year t	Year t+1
Suspected Cheating Classroom			
112A4A342CB214D0001ACD24A3A12DADBCB4A0000000	1.9	5.3	4.4
1B2A34D4AC42D23B141ACD24A3A12DADBCB4A2134141	4.3	5.6	4.3
DB2ABAD1ACBDDA212B1ACD24A3A12DADBCB400000000	3.0	6.5	5.1
1142340C2CBDDAD4B1ACD24A3A12DADBCB43D133BC4	3.6	6.3	4.9
D43A3A24ACB1D32B412ACD24A3A12DADBCB422143BC0	5.2	5.9	4.9
D43AB4D1AC3DD43421245D24X3712DADBCB400000000	4.8	5.3	3.6
DBA2BA21AC3D2AD3C4C4CD40A3A12DADBCB400000000	1.9	6.1	3.6
DBAA4ADC4CBD24DBC2A1110A3A12DADBCB400000000	3.3	6.3	6.2
144A3ADC4CBDDADBCB2C2C2CC43A12DADBCB4211AB343	3.0	6.8	4.9
D43ABA3CACBDDADBCB4A2C2A32L2DADBCB42344B3CB	4.8	7.1	6.6
214AB4DC4CBD31B1B2213C4AD4L2DADBCB4RDB000000	3.6	6.1	4.3
313A3AD1AC3D2A23431223C0000L2DADBCB400000000	3.8	4.7	5.1
D4AAB2124CBDDADBCB1A42CCA34L2DADBCB423134BC1	5.5	6.6	7.7
3B3AB4D14C3D2AD4CBCA1C003A12DADBCB4ADB40000	3.0	6.5	6.6
DBAAB3DCACB1ADADBC42AC2CC310L2DADBCB4ADB40000	3.8	7.1	5.6
DB223A24ACB11A3B24CACD12A241CDADBCB4ADB4B300	4.9	6.5	5.8
D122BA2CACBD1A13211A2D02A2412D0DBCB4ADB4B3C0	3.6	6.1	6.2
1423B4D4A23D24131413234123A243A2413A21441343	4.9	2.5	5.6
DB4ABADCACB1DD3141AC212A3A1C3A144BA2DB41B43	5.9	6.5	7.7
DB2A33DCACBD3D313C21142323CC3000000000000000	3.8	4.4	5.6
1B33B4D4A2B1DADBC3CA22C000000000000000000000	5.0	4.4	7.2
D12443D432D32323C213C22D2C23234C332DB4B300	3.3	3.8	3.6
D4R2341CACBDDAD3142A2344A2AC23421C00ADB4B3CB	6.4	5.9	6.2
	4.1	5.8	5.5
	Average Test Scores		
Typical Classroom			
34AABAD12CBDD3D4C1CA112CAD2CCD000000000000000	3.8	5.6	6.4
D33A3431A2B2D2D44B2ACD2CAD2C2223B400000000000	4.6	4.9	5.8
DB3A431422BD131B4413CD4221A1CDA332342D3AB4C4	4.0	5.1	5.1
D1AA1A11ACB2D3DBC1CA22C23242C3A142B3ADB243C1	4.6	5.9	5.3
D42A12D2A4B1D32B21CA2312A3411D000000000000000	4.5	3.8	6.4
3B2A34344C32D21B1123CDC000000000000000000000	3.3	2.8	5.1
23AA32D2A1BD2431141342C13D212D233C34A3B3B000	3.3	4.4	4.9
D32234D4A1BDD23B242A22C2A1A1CDA2B1BAA33A0000	5.1	5.6	5.9
D3AAB23C4CBDDAD3C322C2A222223232B443B24BC3	4.7	5.6	7.0
D13A14313C31D42B14C421C42332CD2242B3433A3343	2.2	3.8	4.9
D13A3AD122B1DA2B11242DC1A3A121000000000000000	4.5	4.1	5.9
D12A3AD1A13D23D3CB2A21CCADA24D2131B4400000000	3.6	5.3	5.9
314A133C4CBD142141CA424CAD34C122413223BA4B40	3.3	4.7	4.4
D42A3ADCACBDDADBC42AC2C2ADA2CDA341BA3B24321	5.6	6.9	8.5
DBAA34DC2CB2DAD324C412C1ADA2C3A341BA200000000	5.0	5.9	7.0
D1341431ACBDDAD3C4C213412DA22D3D1132A1344B1B	3.8	5.3	5.3
1BA41A21A1B2DAD324CA22C1ADA2CD324132000000000	4.3	5.3	6.8
DBAA33D2A2BDDADBCB4A11C2A2ACCDALB2BA200000000	4.5	6.8	7.9
	4.2	5.1	6.0

Noot: de data in de figuur geven de werkelijke antwoordreeksen weer van twee klassen die dezelfde toets maakten. Elke rij correspondeert met een individuele leerling en een kolom met een volgende vraag. Een letter betekent dat de leerling het juiste antwoord gaf en een nummer betekent een incorrect of een ontbrekend antwoord (0). De totale toetsscores, voor verschillende jaren, staan in de laatste drie kolommen. De bovenste klas wordt verdacht van fraude, de onderste klas niet.

In de Verenigde Staten is gebleken dat scholen in reactie op nieuwe regels, waarbij de bekostiging van scholen af ging hangen van de gemeten leerprestaties, scholen allerlei strategieën zijn gaan volgen waarmee de toetsscores werden verhoogd zonder dat dit samen hoeft te hangen met een

Jacob, B. A. & S.D. Levitt (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. The Quarterly Journal of Economics, 118(3): 843-877.

³² Jacob, B. A. & S.D. Levitt (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. The Quarterly Journal of Economics, 118(3): 843-877.

toename van de vaardigheid die de toets meet: Meer oefenen voor de toets, het verbeteren van de omstandigheden in het toetslokaal, bevorderen dat de leerlingen fit en geconcentreerd op de toets verschijnen, leerlingen niet aan de toets laten deelnemen, het geven van instructies tijdens de toets of het aanpassen van antwoorden na de toets.

2.5 Overzicht toezichtkader van de inspectie

Het toezicht van de inspectie speelt zich af binnen een onderwijssysteem dat gebaseerd is op het principe van vrijheid van onderwijs. Dit betekent dat de oprichting van scholen vrij is, mits aan bepaalde voorwaarden is voldaan en dat ouders vrij zijn een school voor hun kind te kiezen. Een belangrijke consequentie van deze vrijheid van onderwijs is, dat scholen een grote mate van autonomie kennen in hun besluitvorming. Ook betekent het dat de verantwoordelijkheid voor het leveren van goed onderwijs expliciet bij de scholen zelf ligt. Daartoe gebruiken de scholen voornamelijk een systeem van zelfevaluatie, en de inspectie zorgt voor externe controle en evaluatie van de scholen³³. Het toezicht van de inspectie kent een bestuursgerichte aanpak, is risicogericht en op maat, is erop gericht te voorkomen dat scholen (zeer) zwak worden en voorkomen dat voormalig (zeer) zwakke scholen terugvallen in kwaliteit en is gericht op programmatisch handhaven³⁴.

Jaarlijks voert de inspectie voor alle scholen een risicoanalyse uit, op basis van drie soorten gegevens: de opbrengsten van de school, signalen van mogelijke problemen, ook als gevolg van het gevoerde personeelsbeleid, en informatie uit jaarstukken (jaarverslag, jaarrekening, en de zogeheten resultatenbox, die gevuld wordt met kengetallen welke tussen departement en onderwijsveld overeen zijn gekomen). Het doel van de risicoanalyse is te bepalen welke scholen in aanmerking komen voor een basisarrangement en voor welke scholen nader onderzoek nodig is om het toezichtarrangement te kunnen bepalen. Als uit de risicoanalyse blijkt dat bij een school met een basisarrangement de leerresultaten één jaar onder de norm liggen, dan treedt het zogeheten preventief toezicht in werking. Als deze risicoanalyse drie jaar achter elkaar onvoldoende lijkt, dan voert de inspectie een kwaliteitsonderzoek uit om de kwaliteit van de school te beoordelen. Met scholen die na dit kwaliteitsonderzoek (zeer) zwak blijken stelt de inspectie een toezichtplan op, waarin is aangegeven hoe de inspectie het geïntensiveerd toezicht inricht en wat er hierbij van het bestuur wordt verwacht³⁵.

Het kwaliteitsonderzoek van de inspectie kent een viertal onderdelen: (1) oordeel over opbrengsten (onderdeel A van het kernkader toezicht basisonderwijs), (2) oordeel over onderwijsproces (onderdelen B en C van het kernkader toezicht basisonderwijs), (3) oordeel over kwaliteitszorg (onderdeel D van het kernkader toezicht basisonderwijs), en (4) wet- en regelgeving³⁶. Het toezichtkader kent een gelaagde structuur, met een uitgebreide set van 75 indicatoren. Daaruit zijn 45 indicatoren geselecteerd die het kernkader vormen. Daarvan zijn 10 indicatoren voor het primair onderwijs zogeheten normindicatoren (tabel 2.1). Deze spelen een belangrijke rol bij de beslissing of een school de kwalificatie (zeer) zwak krijgt.

³³ Scheerens e.a. (2012). OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes. Enschede: Universiteit Twente.

³⁴ Inspectie van het Onderwijs (2012). Toezichtkader po/vo 2012. Utrecht.

³⁵ Zie hiervoor een schema dat het werkproces van het toezicht duidelijk maakt, in Inspectie van het Onderwijs (2012), pagina 11. Toezichtkader po/vo 2012. Utrecht.

³⁶ Zie bijlage A voor het kernkader toezicht basisonderwijs.

Tabel 2.1 Normindicatoren voor bepaling van toezichtarrangement

Opbrengsten	
1.1	De resultaten van de leerlingen aan het eind van de basisschool liggen ten minste op het niveau dat op grond van de kenmerken van de leerlingenpopulatie mag worden verwacht.
1.2	De resultaten van de leerlingen voor Nederlandse taal en voor rekenen en wiskunde tijdens de schoolperiode liggen ten minste op het niveau dat op grond van de kenmerken van de leerlingenpopulatie mag worden verwacht.
Onderwijsleerproces	
2.1	Bij de aangeboden leerinhouden voor Nederlandse taal en rekenen en wiskunde betreft de school alle kerndoelen als te bereiken doelstellingen.
2.2	De leerinhouden voor Nederlandse taal en voor rekenen en wiskunde worden aan voldoende leerlingen aangeboden tot en met het niveau van leerjaar 8.
2.4	De school met een substantieel aantal leerlingen met een leerlinggewicht biedt bij Nederlandse taal leerinhouden aan die passen bij de onderwijsbehoeften van leerlingen met een taalachterstand.
5.1	De leraren geven duidelijke uitleg van de leerstof.
5.2	De leraren realiseren een taakgerichte werksfeer.
5.3	De leerlingen zijn actief betrokken bij de onderwijsactiviteiten.
Zorg en begeleiding	
7.1	De school gebruikt een samenhangend systeem van genormeerde instrumenten en procedures voor het volgen van de prestaties en de ontwikkeling van de leerlingen.
8.3	De school voert de zorg planmatig uit

Het oordeel op de indicatoren binnen de onderdelen van het kernkader kan de volgende vormen aannemen: slecht, onvoldoende, voldoende, goed, of niet te beoordelen. Het beslisschema komt er grofweg op neer dat als het oordeel over de eindopbrengsten onvoldoende is, er altijd een aangepast toezichtarrangement volgt. Afhankelijk van het aantal onvoldoendes ten aanzien van het onderwijsleerproces en zorg en begeleiding wordt de school zwak of zeer zwak bestempeld. Als het oordeel over de opbrengsten voldoende is, is het toezichtarrangement afhankelijk van het aantal onvoldoendes ten aanzien van het onderwijsleerproces en zorg en begeleiding. Het oordeel over de kwaliteitszorg wordt momenteel alleen in de beoordeling betrokken indien de eindopbrengsten niet te beoordelen zijn. Er wordt nu echter aan een voorstel voor een nieuw toezichtkader gewerkt waarbij juist het oordeel over de kwaliteitszorg explicieter betrokken gaat worden in de beoordeling door de inspectie. Men zou kunnen zeggen dat de zelfevaluatie (hetgeen de kwaliteitszorg in feite betekent) van groter belang gaat worden voor de inspectie.

Naast de normindicatoren die gehanteerd worden voor de beoordeling of een school zwak of zeer zwak is, bevat het toezichtkader ook andere kernindicatoren, zoals die over de doorstroom: “De leerlingen doorlopen in beginsel de school binnen de verwachte periode van 8 jaar”. Opmerkelijk is dat hoewel deze indicatoren zijn opgenomen in het toezichtkader, de inspectie aangeeft dat sommige van deze indicatoren “niet of nauwelijks een rol spelen in het toezicht”.

2.6 Overzicht bekostiging van scholen

Basisscholen ontvangen naast de bekostiging voor het personeel een bekostiging voor de materiële instandhouding (MI-bekostiging). De MI-bekostiging is normatief onderbouwd met programma's van eisen. Ieder programma van eisen omvat een inhoudelijke omschrijving van een van rijkswege

verantwoord geachte voorziening en het bedrag dat hiervoor noodzakelijk is³⁷. Er wordt onderscheid gemaakt tussen groepsafhankelijke programma's van eisen (waaronder onderhoud gebouwen, energie- en waterverbruik en publiekrechtelijke heffingen), leerlingafhankelijke programma's van eisen (waaronder middelen voor medezeggenschap, culturele vorming, dienstreizen etc., en administratie, beheer en bestuur), en aanvullende programma's van eisen (waaronder Nederlands onderwijs aan anderstaligen).

De berekening van de bekostiging is gebaseerd op de volgende formule:

$$Y = Y_a + Y_b + Y_c + Y_d$$

waarin:

- Y = bekostiging per school per jaar
- Y_a = bekostiging groepsafhankelijke programma's van eisen
- Y_b = bekostiging leerlingafhankelijke programma's van eisen
- Y_c = bekostiging aanvullende programma's van eisen
- Y_d = extra bekostiging

Voor elk van de symbolen Y_a tot en met Y_d geldt een formule waarin gerekend wordt met een vast bedrag per school en een bedrag per variabele indicator (leerling, groep of vierkante meter). Het aantal leerlingen van een basisschool wordt bepaald volgens artikel 134 van de Wet op het Primair Onderwijs: het aantal ongewogen leerlingen op 1 oktober van het voorafgaande jaar, verhoogd met 3 procent, waarbij het verkregen aantal naar beneden op een geheel getal wordt afgerond of, als artikel 134, zesde lid van de Wet op het Primair Onderwijs van toepassing is: het feitelijk aantal leerlingen op 1 maart van het jaar waarvoor de bekostiging wordt verstrekt. Het aantal groepen leerlingen wordt langs normatieve weg vastgesteld.

2.7 Hoe kun je strategisch gedrag meten?

Er doet zich een aantal problemen voor bij het meten van strategisch gedrag.

In de eerste plaats zal er – ook zonder strategisch gedrag van scholen – veel variatie zitten in de uitkomsten van scholen. In veel gevallen zal het dus niet duidelijk zijn of een school echt afwijkt van een bepaalde norm, of dat er sprake is van toevalligheden.

In de tweede plaats zal er een neiging bestaan tot beleid waarmee de indicatoren die gebruikt worden voor een beoordeling gunstig worden beïnvloed. Het is daarom onduidelijk of wat gemeten wordt een eerlijke weerspiegeling is van de werkelijkheid of een uiting van het strategische gedrag.

In de derde plaats kan het lastig zijn om een onderscheid te maken tussen gewenst en ongewenst strategische gedrag. Als een school opvallend goed scoort op een bepaalde indicator kan het zijn dat er sprake is van beleid om de waarde van die indicator gunstig te beïnvloeden zonder daarmee de kwaliteit echt te verhogen, maar kan er ook sprake zijn van een werkelijke kwaliteitsimpuls.

Een eerste mogelijkheid is om niet naar de scores van afzonderlijke scholen te kijken, maar naar de verdeling van deze scores over alle scholen. Net zoals bij het voorbeeld van de dobbelsteen van een afzonderlijke deelnemer aan het experiment niet gezegd kan worden of hij het puntental op de dobbelsteen goed doorgeeft of niet, maar dit uit de verdeling wel blijkt, zo kan strategisch gedrag van

³⁷ PO-Raad (2012). Bekostigingsstelsel primair onderwijs: Programma's van eisen voor het jaar 2013. Utrecht.

scholen zichtbaar worden in de verdeling. In de analyse van grootschalige bestanden zullen we van deze aanpak gebruik maken, conform de denkbeeldig voorbeelden eerder in deze paragraaf.

Een andere mogelijkheid om een beeld te krijgen van strategisch gedrag is om gegevens te verzamelen bij scholen die niet benut worden bij het individuele toezicht en waarbij de betrokkenen dus geen reden hebben om niet eerlijk te zijn over het antwoord. Een risico hierbij blijft altijd, dat de ondervraagde niet overtuigd is van de anonimiteit, of een neiging heeft tot sociaal wenselijk antwoorden. Een aanpak om dit te ondervangen is om betrokkenen niet alleen te vragen of zij zelf wel eens bepaald strategische gedrag hebben vertoond, maar ook of ze weten of anderen dat hebben gedaan. Opnieuw kan dan van een individueel antwoord niet worden bepaald of het overeenkomt met de werkelijkheid, maar moet gelden dat de verdeling van de antwoorden voor de eigen school gelijk is aan de verdeling die gevonden wordt voor andere scholen. Verwacht kan worden dat afhankelijk van de rol van de respondenten, de neiging om sociaal wenselijke antwoorden te geven groter zal zijn. Om die reden vergelijken we de vragen aan schoolleiders met de vragen aan docenten.

Informatie kan soms ook worden ingewonnen bij personen die geen belang hebben bij het verborgen houden van strategisch handelen. Om die reden enquêteren we ouders, die voor een aantal onderwerpen een beeld hebben van wat er gebeurt op school. Door hun antwoorden met die van docenten te vergelijken, of door hun antwoorden te vergelijken met officiële cijfers, kan bekeken worden of er sprake is van discrepanties in dit opzicht.

Het voorkomen van gedrag dat door de inspectie wellicht als ongewenst wordt gezien, wil op zich nog niet zeggen dat scholen dit uit strategische overwegingen doen. Het kan ook zijn dat scholen een ander beeld hebben over wat gewenst gedrag is. Als dit gedrag uit strategische redenen wordt gedaan, zullen met name scholen dit doen die hiermee ook hun kwaliteitsoordeel kunnen veranderen. Om die reden zullen we kijken naar het verband tussen het voorkomen van bepaalde vormen van strategisch gedrag en afstand van de opbrengsten van die school ten opzichte van de ondergrens voor de eindtoetsresultaten. Om zicht te krijgen op mogelijke andere motieven, leggen we ook het verband tussen het gedrag van de school en hun kijk op het toezichtkader en het nut daarvan.

Concreet ontstaan hierdoor de volgende indicatoren die een beeld geven over wat er speelt bij bepaalde vormen van ongewenst gedrag:

- Geeft men antwoord op de vraag? Niet antwoorden kan betekenen dat men het niet weet, maar ook dat men liever niet antwoordt.
- Geven respondenten aan dat er sprake is van deze vorm van strategisch gedrag? Een positief antwoord lijkt duidelijk te maken dat er inderdaad sprake is van dit gedrag.
- Wat is het verschil tussen de mate waarin er sprake zou zijn van dit gedrag op de eigen school en dat op andere scholen? Dit verschil laat zien dat er feitelijk meer sprake is van deze vorm van strategisch gedrag maar men dat niet graag toegeeft voor de eigen school.
- Wat is het verschil in antwoord tussen docenten en schoolleiders? Dit verschil wijst er ook op dat een kwestie als gevoelig wordt ervaren, er van uit gaande dat een schoolleider sterker rekening zal houden met gewenste antwoorden dan de docenten.
- Wat is het verschil in antwoord tussen leerkrachten en schoolleiders enerzijds en ouders anderzijds? Als ouders meer dan de scholen aangeven dat bepaalde vormen van gedrag plaatsvinden, kan dat een indicatie zijn dat scholen dit niet toe te willen geven.

- Hoe groot is het verband van verschillende vormen van strategisch gedrag en de afstand van die school tot de ondergrens die de inspectie hanteert? Een dergelijk verband geeft aan in welke mate dit gedrag strategisch wordt ingezet.
- Hoe groot is het verband met de kijk van de school op de opbrengstindicatoren van de inspectie? Een dergelijk verband geeft aan in welke mate er ander redenen zijn voor dit gedrag.

3 Verkenning strategisch gedrag in het primair onderwijs

Voor het onderzoek zijn verkennende gesprekken gevoerd met de inspectie, schoolleiders en leerkrachten van basisscholen, de Algemene Vereniging van Schoolleiders en de PO-Raad. Doel was om het beeld te krijgen van de invloed dat het toezichtkader van de inspectie, en in het bijzonder de opbrengstindicatoren als onderdeel hiervan, heeft op het gedrag van scholen in aanvulling op wat uit de literatuur bekend is. Op basis van deze gesprekken zijn dan de thema's geselecteerd die in de vragenlijst aan de orde komen.

3.1 Gesprekken met schoolleiders en leerkrachten

Op basis van beschikbare data over Nederlandse scholen is een aantal scholen geselecteerd in de regio Nijmegen die op bepaalde kenmerken (waaronder een opmerkelijke stijging van de score op de Cito eindtoets basisonderwijs in de afgelopen jaren, een groei van het aantal instromers van buiten de wijk of een zeer zwakke beoordeling door de inspectie in de laatste jaren) eruit sprongen. Op deze scholen is met de schoolleider en leerkrachten een verkennend gesprek gevoerd op basis van een gestructureerde open vragenlijst. Daarnaast is er een gesprek gevoerd met het AVS waarbij vier schoolleiders en een medewerker van het AVS betrokken waren. Ook is gesproken met twee personen van de PO-Raad. Hieronder volgt een toelichting van de belangrijkste punten die uit deze gesprekken naar voren kwamen.

Over het gebruik van opbrengstindicatoren in het algemeen

Aan de ene kant vindt men dat het werken met het toezichtkader helderheid geeft over waarnaar wordt gekeken. Men is het erover eens dat er indicatoren moeten zijn die richting geven aan het toezicht op onderwijs, of een kader scheppen waarbinnen de scholen zich kunnen bewegen. Er is daarbij waardering voor de manier waarop de inspecteurs zich binnen een kort tijdsbestek voorbereiden en een beeld vormen van wat er op een school gebeurt. Ook kan het helpen om de urgentie van bepaalde zaken bij de eigen leerkrachten onder de aandacht te brengen.

Aan de andere kant vindt men het toezicht een beperkte manier van kijken. Opgemerkt wordt dat het beeld dat ontstaat te eenzijdig is, dat er te weinig rekening wordt gehouden met de context op de betrokken scholen, en dat de indicatoren gericht zijn op traditioneel onderwijs met een leerstofjaarklassensysteem en weinig open staat voor vernieuwing. Er wordt een aantal keer gewezen op het gebruik van het leerlinggewicht als maatstaf voor de sociale achtergrond van de leerlingen. De schoolleiders geven aan dat er qua leerlinggewicht niets aan de hand hoeft te zijn, maar dat er een andere problematiek kan spelen die niet in cijfers tot uitdrukking komt, maar wel een weerslag heeft op de te behalen resultaten. De schoolleiders geven daarbij ook aan dat het niet altijd helemaal duidelijk is wat precies bedoeld wordt met de indicatoren, en een enkeling stoort zich daarbij aan de paternalistische toon van het toezichtkader.

“Het toezichtkader stelt dat ‘scholen een samenhangend systeem moeten hebben’. Maar wat zegt de samenhang van het systeem over de kwaliteit ervan? Ook ‘moeten scholen moeten vroegtijdig de zorgbehoefte bij leerlingen signaleren’. Maar dit zegt niks over of je de goede zorg biedt aan een leerling die dat nodig heeft.”

Een aantal schoolleiders deelt de mening dat met de focus op het meetbare en de huidige indicatoren de vakbekwaamheid van leerkrachten wordt

uitgekleed en er een angstcultuur ontstaat. De keuze voor meetbare zaken is een logische, maar het dekt de lading niet. Onderwijs is meer dan alleen toetsen. De PO-Raad ziet hier een belangrijke rol weggelegd voor de schoolbesturen. Zij staan dicht bij de scholen dan de inspectie en zien dus beter wat er feitelijk gebeurt, maar kunnen toch ook eenzelfde kwaliteitsbevorderende rol hebben als de inspectie.

Over de eindtoets basisonderwijs

Alle schoolleiders geven aan dat de keuze voor het type eindtoets door het schoolbestuur wordt gemaakt. Daarbij speelt gemak een belangrijke rol, men heeft “geen zin in gedoe met de inspectie”. Toch merken ze ook op dat als je een andere toets hebt, het ook lastiger is om je opbrengsten te verantwoorden en te vergelijken met andere, concurrerende scholen. Er wordt in de gesprekken op gewezen dat er momenteel een te sterke nadruk ligt op de resultaten van de toetsen, zowel die van de eindtoets basisonderwijs, als die van de toetsen in het leerlingvolgsysteem.

Men vindt dat de gehanteerde toetsen een te beperkt beeld geven van wat een leerling kan. De scores meten kennis in rekenen en taal, maar zeggen niets over de betrokkenheid van kinderen, hun creatief denkvermogen, hun oplossingsgerichtheid et cetera. Juist de competenties die niet cognitief zijn, kunnen maken dat een kind later een goeie timmerman of dierenverzorger kan worden. Ook heeft men soms het gevoel dat de leerkrachten zich zelf ook teveel beoordeeld voelen door de nadruk op de toetsresultaten. Ook vindt men dat er te weinig verschillende toetsen beschikbaar zijn. Nu is alles gerelateerd aan Cito. Dit maakt dat scholen “Cito-gerelateerd onderwijs” gaan geven.

“We hebben nu een taalzwak kind dat heel laag op de taaltoetsen uit het leerlingvolgsysteem scoort, tenzij je juist de specifieke woorden uit die toets gaat oefenen met dat kind. Maar dat is de verkeerde manier om dergelijke toetsen te gebruiken. Dit wordt wel uitgelokt door de afrekencultuur van de inspectie. Het zou beter zijn om met de toetsen te kijken wat de groei van een kind is en waarom het kind taalzwak is, en wat hiervoor een oplossing is.”

“Soms moet ik tegen een leerkracht zeggen dat zijn leerlingen juist niet alles goed moeten hebben bij een toets uit het leerlingvolgsysteem. Een goede toets meet juist waar een kind staat, dus moeten er juist onderdelen in zitten die de meeste kinderen (nog) niet weten. Maar de leerkracht voelt zichzelf ook beoordeeld op basis van die toetsresultaten. Het leidt allemaal af van waar het echt om gaat: naar het individuele kind kijken.”

“Er zijn zeker scholen die dan eens goed naar de toets gaan kijken en nog gaan oefenen op onderdelen uit de toets. Daar word je toe aangezet, vanwege het oordeel van de inspectie. En dan zijn er ook nog allerlei commerciële partijen die zich op het trainen van de eindtoets richten en gaat het niet alleen om het oefenen van de eindtoets, maar ook om eerdere toetsen in het leerlingvolgsysteem in de groepen 5, 6 en 7. Daar ga je als school toch in mee. Dus ‘je leert het hondje een trucje zodat ze het goed doet voor het baasje’.”

“Waarom zou je niet alleen einddoelen stellen en scholen de vrijheid geven om te bepalen hoe en wanneer ze daar komen. Dat zou zoveel druk wegnemen. En je zorgt dat echt goede scholen komen bovendrijven. Nu heet je een goede school als je gemiddeld 540 op de eindtoets scoort. Maar hoeveel is er geoefend? Hebben ze het zelf wel gemaakt? Er zijn zoveel fraudemogelijkheden. Of die school die een hele hoge uitstroom naar het voortgezet onderwijs had, hoeveel procent van die leerlingen stroomt later nog af door een te hoge toetsscore en daarbijbehorend advies?”

Deze opmerkingen van de schoolleiders over de gerichtheid op de Cito-toetsen, sluit aan bij het in het theoretisch deel gegeven voorbeeld van strategisch gedrag, waarbij de indicatoren de scholen stimuleren om meer aandacht te besteden aan de gemeten indicatoren en minder aan de niet gemeten indicatoren. De schoolleiders zijn daarom gevraagd naar hun ervaring met oefenen voor de eindtoets basisonderwijs, door eigen leerkrachten en meer in het algemeen door ouders buiten de school om. Het merendeel van de schoolleiders zegt dat er op hun school geen bewust beleid is met het oefenen van toetsen, maar dat men zich kan voorstellen dat het gebeurt. De schoolleiders maken zich wel zorgen over het oefenen met de toetsen en de commerciële training die er buitenschools wordt

aangeboden. Als steeds meer scholen gaan oefenen, gaan de gemiddelde scores omhoog en gaan de normscores omhoog. Als je dan als school niet oefent, blijf je achter. De schoolleiders zouden daarom minder gerichtheid op specifieke toetsen wensen.

Over zittenblijven, doorstroom of instroom

Over het algemeen zegt men dat men zich bij de beslissing of een leerling al dan niet moet blijven zitten, niet afvraagt wat dit voor gevolgen heeft voor de beoordeling door de inspectie. De inspectie geeft ook aan zelden naar deze indicator te kijken bij de beoordeling van de kwaliteit van het onderwijs op een school. Slechts één schoolleider gaf te kennen, dat hij wilde voorkomen dat hij “gezeur met de inspectie” kreeg. Maar als het nodig is, blijft een kind zitten of wordt de kleuterperiode verlengd. Per kind wordt gekeken of het toe is aan de volgende stap, niet alleen bij de overgang van groep 2 naar groep 3, maar bij elke jaarovergang. “Ik hou daarbij geen rekening met wat volgens de inspectie zou moeten, dan ben ik liever eigenwijs.”

Wat betreft instroom, is het beeld dat schoolleiders geven iets gemengder dan met betrekking tot het zittenblijven. Men zegt over de instroom niet keihard dat er rekening gehouden moet worden met het toezichtkader van de inspectie, maar men geeft wel te kennen dat het op de achtergrond een rol speelt. Meer dan bij het zittenblijven. Zo kent één van de schoolleiders waarmee gesproken is wel een school die alleen kinderen aanneemt van ouders die minimaal een havo-diploma hebben. En dat men deze verleiding kan begrijpen omdat er zo'n sterke gerichtheid is op de toetsscores.

“Er is laatst een jongen in groep 7 ingestroomd die rekenproblemen had op de vorige school. Moet je die dan weigeren omdat je wordt afgerekend op je zwakke leerlingen? Nee. Maar je past toch enige calculatie toe, want er is meteen het besef dat deze jongen niet mee gaat tellen met de score op de Eindtoets Basisonderwijs (omdat hij pas in groep 7 is ingestroomd).”

De rol van de inspecteur: van politie-agent naar adviseur/coach

Alle schoolleiders geven aan dat ze de inspecteur van grote waarde vinden, maar dat deze teveel in de controlerende taak zit. Ze zeggen dat er wel grote verschillen zijn tussen de inspecteurs, en dat sommige tussen de regels door wel advies geven, maar dat dit meestal niet expliciet is. Dat dit ook niet bij hun huidige rol als controleur past. Voor hun gevoel is dit in de afgelopen tien jaar steeds meer het geval geworden en zijn opbrengsten belangrijker geworden dan de processen. Men erkent het belang van signaleren of de indicatoren gehaald zijn, maar men mist het begrip voor het proces erachter, of het sociale aspect.

“De eindtoets basisonderwijs is dit jaar onverwacht slecht gemaakt op onze school. We zoeken nu naar de oorzaken van die lagere resultaten. Dat is heel complex. Dat kan ik wel uitleggen aan de inspecteur, maar daar wordt uiteindelijk in het oordeel niks mee gedaan.”

Men ziet een grotere waarde van een inspecteur in de functie van een sparringpartner of adviseur. Minder protocollen, meer gesprek en meer vertrouwen. Een inspectie die meekijkt en adviseert over waar de school

“De inspectie zou breder moeten kijken en meer moeten adviseren, en hun oordeel niet baseren op een bezoek van een halve of hele dag. Dat geldt trouwens ook voor de eindtoets. In drie dagen wordt bepaald wat de opbrengst van acht jaar onderwijs is. Het is een circus geworden en er ligt zoveel druk op.”

aan moet werken, meer partnerschap. Bij enkele besturen, en ook door AVS, wordt contact gezocht met inspecteurs buiten het officiële inspectiemoment om. Ze worden bijvoorbeeld uitgenodigd voor studiemiddagen, of een training over waar de school op moet letten in het kader van de inspectie. Zo hoopt men enerzijds de inspectie wat mee te kunnen nemen in bepaalde ontwikkelingen en keuzes en

anderzijds de ‘angst’ voor de inspectie weg te nemen. Hier wordt door enkele schoolleiders overigens wel gewezen op de eigen verantwoordelijkheid: “*Een goede directeur gaat het gesprek met de inspecteur aan in de zoektocht naar verbetering*”. Men spreekt over een omslag in het denken en omgaan met de inspectie, van een top-down benadering vanuit de inspectie naar een bottom-up benadering vanuit de scholen.

3.2 Gesprek met de inspectie

Naast gesprekken met schoolleiders is er een gesprek gevoerd met een aantal inspecteurs. Deze wijzen voornamelijk op strategisch gedrag met betrekking tot de eindresultaten. Er zijn verschillende manieren waarop scholen deze opbrengstindicator gunstig kunnen beïnvloeden. Het kan gaan om gedrag waarmee de gemiddelde schoolprestaties hoger lijken dan ze werkelijk zijn, afwijken van de afnamevoorschriften, of een bovenmatige aandacht voor de toetsen.

Bij het gunstiger laten lijken van de scores kan gedacht worden aan het uitsluiten van leerlingen, of het uit laten stromen van leerlingen in groep 7. Ook kan men leerlingen laten herkansen als het eerste resultaat tegenvalt of aan meer dan één toets mee te laten doen (bijvoorbeeld B8 en M8 of drempeonderzoek en Cito eindtoets). Afwijken van afnamevoorschriften kan voorkomen in allerlei gradaties variërend van het voorlezen van vragen, het helpen bij vragen, leerlingen hulpmiddelen laten gebruiken die niet zijn toegestaan (rekenmachine), of leerlingen die niet dyslectisch zijn toch deze toets laten maken, tot het achteraf aanpassen van antwoorden. Bovenmatige aandacht voor de toets kan bijvoorbeeld door leerlingen op school of thuis veel te laten oefenen, de lesinhoud af te stemmen op de inhoud van de toets, of het verschuiven van de schooltijd reken- en taalonderwijs.

De inspectie is op zoek naar mogelijkheden om deze vormen van strategisch gedrag te constateren, bijvoorbeeld door gegevens te verzamelen over meervoudig afnemen van toetsen, deelname aan de toets te analyseren en opmerkelijke uitslagen op te sporen met statistische analyses. Het probleem hierbij is dat dergelijke analyses statistisch gezien niet krachtig genoeg zijn om met voldoende grote betrouwbaarheid over een individuele school aan te kunnen geven of er sprake is van afwijkende uitslagen. Een opvallende stijging van een score bij een school hoeft geen manipulatie te zijn, maar kan ook het gevolg zijn van een daadwerkelijke verbetering van de schoolkwaliteit, of een verandering van de leerlingpopulatie.

3.3 Keuze van thema's

Op basis van de literatuur en de gesprekken met schoolleiders, leerkrachten en inspecteurs kan afgewogen worden welke aspecten van het schoolbeleid het meest gevoelig zijn voor strategisch gedrag. Daarbij wordt het meest gewezen in de richting van de eindopbrengsten. Hierbij kan sprake zijn van het buiten de toets houden van leerlingen, *teaching to the test*, buitensporig oefenen in de periode voor de toets en het beïnvloeden van de afname. Om die reden hebben we een groot aantal vragen hierover opgenomen in de vragenlijst voor schoolleiders en docenten en enkele in de vragenlijst voor ouders.

Daarnaast wordt in de gesprekken af en toe gewezen op strategisch gedrag met betrekking tot de doorloop snelheid van leerlingen en de advisering voor het voortgezet onderwijs. Daarom hebben wij

ook hierover enkele vragen in de vragenlijsten opgenomen. De doorloopsnelheid is daarnaast ook om een aantal andere redenen interessant om te bestuderen. De indicator is zeer beïnvloedbaar, maar ook congruent: Het is moeilijk voor te stellen dat de indicator niet overeenkomt met de werkelijkheid. Daarnaast heeft de doorloopsnelheid van leerlingen ook consequenties voor de bekostiging van scholen. Scholen hebben dus veel belang bij deze indicator, kunnen hem goed strategisch beïnvloeden, maar kunnen er ook voor zorgen dat de indicator voldoet aan de eisen van de inspectie. Daardoor kan het zijn dat hoewel de indicator (of de bekostigingsregel) aanleiding is voor strategisch gedrag, er toch weinig feitelijke onvoldoende opbrengsten op deze indicator worden geobserveerd. Deels omdat deze volgens de inspectie ook zelden wordt meegenomen in de beoordeling. Dat is consistent met de observatie dat in het veld dit aspect meer genoemd wordt dan bij de inspectie. De doorloopsnelheid leent zich ook bij uitstek om te illustreren hoe in data-analyse aanwijzingen voor strategisch gedrag kunnen worden gevonden.

De indicator met betrekking tot de ontwikkeling van leerlingen met specifieke onderwijsbehoeften wordt niet genoemd. Een reden hiervoor kan zijn dat de beoordeling hiervan vrij subjectief is en niet makkelijk manipuleerbaar is voor scholen. Ook sociale competenties worden in de gesprekken die we hebben gevoerd niet naar voren gehaald. We hebben daarom over deze aspecten geen vragen gesteld.

Een meer algemene vorm van strategisch gedrag is het niet toelaten van leerlingen die op één of meerdere indicatoren voor negatieve opbrengsten kunnen zorgen. We hebben daarom zowel in de ouder- als de leerkracht/schoolleidersvragenlijsten vragen opgenomen over de toelating van potentieel zwakke leerlingen en de doorverwijzing van leerlingen naar het speciaal onderwijs.

Omdat vanwege het strategisch gedrag dat plaats kan vinden rondom de eindtoets basisonderwijs, door de inspectie een vergelijking kan worden gemaakt met uitslagen van eerdere toetsen, vragen we in de vragenlijst voor leerkrachten en schoolleiders ook naar het gedrag met betrekking tot deze toetsen.

De andere onderdelen van het toezichtkader worden niet of nauwelijks genoemd in de literatuur of de gesprekken. In principe zouden vragenlijsten met betrekking tot veiligheid op eenzelfde wijze beïnvloed kunnen worden als toetscijfers, maar er zijn geen aanwijzingen dat hier sprake van lijkt te zijn. Indicatoren die vastgesteld worden op basis van een bezoek door een inspecteur, zullen uiteraard ook enigszins gemanipuleerd worden doordat scholen hun beste beentje voor zullen zetten op de dag dat de inspecteur langskomt. Voor een deel zal de inspecteur dit echter kunnen doorzien, terwijl er bovendien geen aanwijzingen zijn dat de neiging om een mooie indruk te maken, sterk verschilt tussen scholen. Zelfevaluatie als inspectiecriteria is nog in ontwikkeling. Op den duur zal hier ook rekening gehouden moeten worden met strategisch gedrag, maar momenteel zijn er nog geen aanwijzingen dat dit al speelt.

4 Analyses van schoolpraktijk op basis van vragenlijsten

In dit hoofdstuk beschrijven we de resultaten van analyses op basis van gegevens verzameld onder schoolleiders, leerkrachten en ouders over de opbrengstindicatoren van de inspectie. De data uit de vragenlijsten wordt daarbij gerelateerd aan informatie over hoe ver de school zich van de door de inspectie vastgestelde ondergrens voor de Cito eindtoets bevindt, en enkele schoolkenmerken waaronder regio, schoolgrootte en denominatie.

4.1 Schoolleiders en leerkrachten

4.1.1 Aanpak en respons schoolleiders en leerkrachten

Er zijn in vier fasen uitnodigingsbrieven voor het onderzoek verstuurd naar schoolleiders en leerkrachten. In eerste instantie zijn naar 1.000 scholen brieven verstuurd waarin werd aangekondigd dat de scholen telefonisch benaderd zouden worden. Vrij snel na het starten van de telefonische enquête bleek dat het nauwelijks mogelijk was om leerkrachten telefonisch te bevragen. Om de respons onder met name de leerkrachten te verhogen is vervolgens voor twee aanvullende benaderingswijzen gekozen. Er zijn in drie tranches in totaal 2.497 brieven verstuurd met daarin een schriftelijke vragenlijst voor de schoolleider, één voor leerkrachten van groep 3 tot en met 7 en een voor leerkrachten van groep 8. Daarnaast is in de telefonische enquêtes met de schoolleiders gevraagd naar e-mailadressen zodat we leerkrachten via een e-mail konden uitnodigen voor een webenquête. Hiermee zijn 1.302 e-mailadressen verkregen. Tabel 4.1 laat de respons zien die op basis van deze aanpak is behaald en tabel 4.2 laat de redenen zien die de respondenten aangaven bij de telefonische enquête om niet mee te doen aan het onderzoek.

Tabel 4.1 Respons onder schoolleiders en leerkrachten (absolute aantallen)

	Schoolleiders	Groep 3-7	Groep 8	Totaal
Telefonisch	480	58	27	565
Schriftelijk	388	573	531	1492
Web		214	191	405
Totaal	868	845	749	2462

Tabel 4.2 Redenen om niet mee te doen in telefonische enquête

Reden:	Absoluut aantal	Percentage
Kost teveel tijd/moeite	23	21.50
Te vaak enquêtes	18	16.82
Heeft geen tijd/bezig	18	16.82
Geen opgaaf van reden	13	12.15
Principieel tegen enquêtes	9	8.41
Geen interesse in onderwerp	7	6.54
Anders	19	17.76
Totaal	107	100

Ook hebben we de antwoorden van de schoolleiders en leerkrachten gerelateerd aan informatie over hoe ver de school zich van de door de inspectie vastgestelde ondergrens voor de Cito eindtoets

bevindt en enkele schoolkenmerken waaronder regio, schoolgrootte en denominatie. Hiervoor koppelen we het voor dit onderzoek verzamelde databestand met de vragen voor schoolleiders en leerkrachten aan een databestand met gegevens over gemiddelde toetsscores en het percentage gewichtenleerlingen van de inspectie. Van de 2.462 respondenten in het door ons verzamelde bestand zijn er 2.419 gekoppeld. Voor 43 respondenten bleek de koppeling niet mogelijk omdat het brinnummer ontbreekt. Uit dit bestand met schoolkenmerken gebruiken we de volgende variabelen:

- Afstand tot de ondergrens³⁸. Hoe ver zit de school met de score op de Cito eindtoets van de ondergrens af? In het toezichtkader van de inspectie staat aangegeven hoe deze ondergrens afhangt van het percentage gewichtenleerlingen op de school. Leerlingen die voldoen aan bepaalde criteria worden hierbij buiten beschouwing gelaten. De inspectie corrigeert soms de gegevens zoals ze die ontvangt. Wij gebruiken deze gecorrigeerde score. We nemen het gemiddelde van de toetsscores in de schooljaren 2011/12-2013/14. Een positieve afstand betekent dat de score van de school hoger is dan de ondergrens.
- Schoolgrootte. Als proxy voor de grootte van de school nemen we het aantal leerlingen dat de Cito eindtoets heeft gemaakt. We nemen het gemiddelde over de schooljaren 2011/12-2013/14.
- Provincie. We nemen de provincie waarin de school ligt. Daarnaast hebben we in andere modellen een dummy meegenomen voor de provincies Groningen, Friesland en Flevoland, omdat deze een groot percentage zwakke scholen kennen.
- Denominatie. We onderscheiden de volgende categorieën:
 - Openbaar
 - Protestants-Christelijk / Reformatorisch
 - Rooms-Katholiek
 - Overige religies: Islamitisch / Hindoeïstisch / Interconfessioneel
 - Overige: Algemeen Bijzonder / Antroposofisch / Gereformeerd Vrijgemaakt

4.1.2 Bekendheid met het toezichtkader van de inspectie

Aan de schoolleiders en leerkrachten is eerst gevraagd in hoeverre men bekend is met het toezichtkader van de inspectie en welke invloed men denkt dat hiervan uitgaat op het onderwijs. Tabel 4.3 laat zien dat een kleine 80 procent van de respondenten zegt het toezichtkader en de criteria goed te kennen. Bijna een vijfde van de respondenten geeft aan dit niet zo goed, of helemaal niet, te kennen. De tabel laat ook zien dat bijna 80 procent van de ondervraagde schoolleiders en leerkrachten het goed of heel goed idee vindt dat de inspectie met opbrengstindicatoren de kwaliteit van scholen in kaart probeert te brengen, al noemt men vaak aanvullend wat kritische kanttekeningen met name ten aanzien van de eenzijdigheid ervan.

“De indicatoren zijn op zich goed. wat ik wel lastig vind is dat de inspectie als het echt landelijk goed gaat de indicatoren opschroeft. Alles wat gemeten wordt, zijn zaken die duidelijk zijn, maar er zijn ook genoeg zaken die wel invloed hebben op de kwaliteit die nu niet gemeten worden en lastiger te meten zijn.”

“Het gebruik van indicatoren is een goede zaak, maar er kan meer rekening gehouden worden met omstandigheden die de resultaten beïnvloeden. Kijk niet alleen naar de opbrengsten, kinderen zijn zoveel meer!”

³⁸ Voor 2.414 van de 2.462 scholen in de data is bekend welk arrangement ze in 2013 hadden. Van de scholen in de data hebben negen scholen een aangepast zeer zwak arrangement en 42 een aangepast arrangement. Deze aantallen zijn erg klein en besloten is om deze variabele niet mee te nemen.

Tabel 4.3 Bekendheid met het toezichtkader van de inspectie en gebruik opbrengstindicatoren voor beoordeling kwaliteit van scholen

Vindt u zelf dat u goed op de hoogte bent van de inhoud van het toezichtkader en de indicatoren van de inspectie?	
Helemaal niet goed	1.83
Niet zo goed	17.26
Goed	67.99
Heel goed	11.21
Antwoord ontbreekt	1.71
Vindt u het, in het algemeen, een goede zaak dat de inspectie met opbrengstindicatoren werkt om de kwaliteit van scholen in kaart probeert te brengen?	
Heel slecht	1.95
Slecht	13.57
Goed	71.73
Heel goed	7.11
Geen mening	2.52
Antwoord ontbreekt	3.13

De tabel laat percentages van het totaal aantal respondenten (2.462) zien.

Vervolgens is gevraagd of men vindt dat de gehanteerde criteria voor het toezicht door de inspectie de juiste zijn. Tabel 4.4 laat zien dat zo'n 44 procent van de ondervraagde schoolleiders en leerkrachten vindt dat de gehanteerde criteria grotendeels de juiste zijn. Dit betekent niet dat men niet kritisch is ten aanzien van deze criteria. Een ongeveer even groot deel van de respondenten denkt dat de indicatoren slechts enigszins de juiste zijn en een kleine zeven procent vindt dat de criteria helemaal niet de juiste zijn.

“Cito neemt steeds vaker de overhand bij het toezicht van de inspectie en bij de doorverwijzing naar het voortgezet onderwijs, maar ook andere indicatoren zijn van belang.”

“Het is jammer dat de inspectie zich zo richt op rekenen en taal. Wij vinden ook samenwerken, zelfregulerend leren, zelfstandigheid, en andere creatieve zaken belangrijk. En daar wordt totaal geen rekening mee gehouden. Wat dat betreft is er een afrekencultuur van de zogenaamde meetbaarheid van onderwijskwaliteit. Er wordt bijvoorbeeld door de inspecteur geopperd om muziekles te laten vervallen, dit is absoluut een vershraling van onderwijs.”

“Er is een verschuiving gaande van harde opbrengsten naar het monitoren van het leerproces/totstandkoming van de opbrengsten. En dat is positief.”

Tabel 4.4 Opbrengstindicatoren zijn de juiste indicatoren?

Vindt u dat de opbrengstindicatoren die worden gehanteerd door de inspectie de juiste indicatoren zijn?	
Helemaal niet	6.38
Enigszins	41.75
Grotendeels	42.20
Ja helemaal	2.52
Geen mening	2.80
Antwoord ontbreekt	4.35

De tabel laat percentages van het totaal aantal respondenten (2.462) zien.

Tot slot is aan de respondenten gevraagd in hoeverre ze denken dat het gebruik van opbrengst-indicatoren een positief of negatief effect heeft op het onderwijs op school. Tabel 4.5 toont dat ruim een derde van de ondervraagden vindt dat het gebruik een negatief effect heeft op het onderwijs.

Tabel 4.5 Effect van gebruik opbrengstindicatoren op onderwijs op school

Vindt u dat het gebruik van opbrengstindicatoren door de inspectie een positief effect heeft op het onderwijs op de school of vindt u het effect er van juist negatief?	
Heel positief	6.70
Een beetje positief	49.80
Vrij negatief	22.42
Negatief	12.14
Weet ik niet	4.14
Antwoord ontbreekt	4.79

De tabel laat percentages van het totaal aantal respondenten (2462) zien.

Tabel 4.6 laat zien in hoeverre deze vragen met elkaar samenhangen en hoe de antwoorden zich verhouden tot hoe ver scholen zich van de ondergrens voor de Cito eindtoets bevinden. Er is een positieve samenhang tussen de diverse stellingen. Respondenten die vinden dat het gebruik van indicatoren een goede zaak is, vinden vaker dat de gehanteerde indicatoren de juiste zijn of dat het gebruik ervan positieve effecten heeft op het onderwijs. Ook is er een positieve relatie met de afstand tot de ondergrens: scholen die beter presteren zijn vaker positief over het gebruik van opbrengstindicatoren in het algemeen en achten de gehanteerde indicatoren vaker de juiste. Met de overige schoolkenmerken wordt geen systematische relatie gevonden. Bij Rooms-Katholieke en scholen op basis van overige religies is de positieve relatie tussen afstand en of men de indicatoren juist vindt sterker.

Tabel 4.6 Onderlinge samenhang en relatie met schoolkenmerken zoals afstand tot ondergrens

Gebruik van indicatoren is in algemeen een goede zaak			
	Model 1	Model 2	
Bekend met toezichtkader	0.09*** (0.019)	0.08*** (0.021)	
Afstand tot ondergrens		0.01** (0.005)	
Constante	2.64 (0.056)	2.64 (0.064)	
Aantal observaties	2322	1722	
R ²	0.0089	0.0113	
De gebruikte indicatoren zijn de juiste			
	Model 1	Model 2	
Afstand tot ondergrens	0.03*** (0.006)	0.02*** (0.006)	
Gebruik van indicatoren is in het algemeen geen goede zaak		0.46*** (0.028)	
Constante	2.43 (0.018)	1.10 (0.083)	
Aantal observaties	1695	1638	
R ²	0.0111	0.1502	

Het gebruik van opbrengstindicatoren heeft positieve effecten op het onderwijs			
	Model 1	Model 2	Model 3
Afstand tot ondergrens	0.02*** (0.008)	0.02** (0.007)	0.01 (0.006)
Gebruik van indicatoren is in het algemeen geen goede zaak		0.68*** (0.033)	0.47*** (0.034)
Gebruikte indicatoren zijn de juiste			0.45*** (0.028)
Constante	2.54 (0.023)	0.60 (0.096)	0.121 (0.096)
Aantal observaties	1659	1607	1540
R ²	0.0058	0.2155	0.3294

De tabel toont de coëfficiënten van een regressiemodel, met de standaardfouten tussen haakjes.

4.1.3 *Schoolleiders en leerkrachten over de rol van de opbrengstindicatoren bij toelating, zittenblijven en doorverwijzing*

De schoolleiders zijn gevraagd in hoeverre de opbrengstindicatoren een rol spelen bij de toelating van leerlingen van wie het vermoeden bestaat dat het zwakke leerlingen zijn. Er is eerst gevraagd in hoeverre men denkt dat dit op de eigen school gebeurt, en vervolgens in hoeverre dit op andere scholen gebeurt. Tabel 4.7 laat zien dat ongeveer vijf procent van de ondervraagde schoolleiders zegt dat dit vrij of heel sterk op de eigen school gebeurt. Daarentegen denkt ongeveer een vijfde dat dit vrij of heel sterk op andere scholen gebeurt. Ook zegt bijna 80 procent van de schoolleiders dat het helemaal niet op de eigen school gebeurt, terwijl dit 20 procent is als het gaat om wat er op andere scholen gebeurt.

Tabel 4.7 Rol opbrengstindicatoren bij beslissing over toelating mogelijk zwakke leerling

Stel dat er een nieuwe leerling wordt aangemeld voor groep 1 van wie wordt vermoed dat het een zwakke leerling zal zijn. In hoeverre wegen bij de beslissing om deze leerling al dan niet toe te laten de consequenties voor de opbrengstbeoordeling door de inspectie mee?		
	Op de eigen school	Op andere scholen
Helemaal niet	78.80	19.70
Een beetje	16.47	30.30
Vrij sterk	3.11	14.98
Heel sterk	1.27	5.99
Weet ik niet	0.00	28.11
Antwoord ontbreekt	0.35	0.92

De tabel laat percentages van het totaal aantal respondenten (868, alleen schoolleiders) zien.

Vervolgens is aan de schoolleiders en aan leerkrachten die lesgeven in de groepen drie tot en met zeven gevraagd in hoeverre de beoordeling door de inspectie een rol speelt bij de beslissing of een leerling moet blijven zitten of niet. Tabel 4.8 laat zien dat de opbrengstindicatoren van de inspectie bij de beslissing over zittenblijven een iets grotere rol spelen dan bij de toelating van een mogelijk zwakke leerling. Dit is opmerkelijk omdat de betreffende indicator volgens de inspectie geen enkele rol speelt bij het toezicht. Opnieuw zien we ook hier dat men denkt dat het bij andere scholen wel speelt, al geeft hier bijna de helft als antwoord “weet ik niet”. De respondenten die “ik weet het niet”

antwoorden als het om andere scholen gaat, blijken in 63 procent van de gevallen “helemaal niet” op de eigen school te antwoorden en in 23 procent “een beetje”.

Tabel 4.8 Rol opbrengstindicatoren bij beslissing over zittenblijven

Als er op uw school beslist moet worden of een leerling doorgaat naar een hogere groep of een jaar overdoet, in hoeverre speelt daarbij dan de beoordeling door de inspectie een rol?		
	Op de eigen school	Op andere scholen
Helemaal niet	64.22	16.45
Een beetje	22.68	23.14
Vrij sterk	7.10	11.37
Heel sterk	1.10	2.60
Weet ik niet	3.00	44.20
Antwoord ontbreekt	1.90	2.25

De tabel laat percentages van het totaal aantal respondenten (1.733, alleen schoolleiders en leerkrachten groep 3-7) zien.

Ook is, aan zowel schoolleiders als alle leerkrachten, gevraagd in hoeverre de beoordeling door de inspectie een rol speelt bij de mogelijke doorverwijzing van leerlingen naar het speciaal basisonderwijs. Tabel 4.9 toont dat iets minder dan tien procent zegt dat de opbrengstindicatoren een vrij of heel sterke rol spelen op de eigen school bij de beslissing of een leerling wordt doorverwezen naar het speciaal onderwijs. Het percentage “speelt vrij tot heel sterk een rol” is iets hoger als het gaat om wat men denkt van andere scholen. Van de respondenten die bij andere scholen “ik weet het niet” invullen, zegt 20 procent dat de indicatoren een beetje meespelen bij de beslissing op de eigen school, en acht procent vrij of heel sterk.

“Je moet voldoen aan inspectienormen, al gaat dat soms tegen je gevoel in”.

“Als ze het reguliere onderwijs echt niet aankunnen en sociaal gezien zich niet prettig voelen, heb je geen andere keus.”

Tabel 4.9 Rol opbrengstindicatoren bij beslissing doorverwijzing naar speciaal onderwijs

In hoeverre speelt op uw school bij het al dan niet doorverwijzen van leerlingen naar het speciaal onderwijs de beoordeling van de inspectie een rol?		
	Op de eigen school	Op andere scholen
Helemaal niet	62.63	13.12
Een beetje	17.42	16.61
Vrij sterk	6.38	9.75
Heel sterk	2.15	2.84
Weet ik niet	7.84	53.74
Antwoord ontbreekt	3.57	3.94

De tabel laat percentages van het totaal aantal respondenten (2.462) zien.

Ook is gevraagd of er vaker of minder vaak doorverwijzingen zouden zijn als de schoolresultaten geen rol zouden spelen bij de beoordeling door de inspectie. Tabel 4.10 laat zien dat het merendeel (bijna 90 procent) aangeeft dat dit niet anders zou zijn. Toch zegt vier procent minder vaak door te verwijzen.

“Ik zou meer kinderen binnen onze school willen houden, maar dan kom je in de knoop met de beoordeling. Hierdoor nemen wij zwakke zij-instromers ook minder snel aan. Kunnen het ons niet meer permitteren, ook al zouden wij het graag willen.”

Tabel 4.10 Meer of minder doorverwijzing naar speciaal onderwijs als geen rol in oordeel

Als de schoolresultaten geen rol zouden spelen bij de beoordeling van de inspectie zou u dan vaker of juist minder vaak leerlingen naar het speciaal onderwijs doorverwijzen?	
Vaker	2.48
Hetzelfde als nu	88.99
Minder vaak	4.31
Antwoord ontbreekt	4.22

De tabel laat percentages van het totaal aantal respondenten (2.462) zien.

Tabel 4.11 laat de resultaten zien van regressiemodellen waarbij de antwoorden op deze vragen zijn gerelateerd aan schoolkenmerken en aan hoe men tegen het gebruik van opbrengstindicatoren in het onderwijs aankijkt. De resultaten laten vooral zien dat hoe minder men denkt dat het gebruik van indicatoren in het onderwijs een goede zaak is of dat het de gehanteerde indicatoren de juiste zijn, hoe vaker men zegt dat de opbrengstindicatoren op *andere* scholen een rol spelen bij toelating, zittenblijven of doorverwijzen naar speciaal onderwijs. Dit verband is niet aanwezig als het over de eigen school gaat. Voor andere schoolkenmerken, zoals grootte, regio of denominatie hangen niet systematisch samen met deze antwoorden.

Tabel 4.11 Regressieresultaten: indicatoren van inspectie spelen een rol bij toelating mogelijk zwakke leerling (alleen schoolleiders)

Rol opbrengstindicatoren bij toelating mogelijk zwakke leerling						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.01 (0.008)	-0.01 (0.008)	-0.01 (0.008)	-0.02 (0.015)	-0.02 (0.015)	-0.01 (0.016)
Gebruik indicatoren in algemeen goede zaak		-0.12*** (0.05)	-0.11** (0.05)		-0.23*** (0.084)	-0.14 (0.089)
Gebruikte indicatoren zijn de juiste			-0.04 (0.036)			-0.23*** (0.069)
Constante	1.272 (0.026)	1.63 (0.138)	1.70 (0.146)	2.08 (0.046)	2.74 (0.254)	3.06 (0.270)
Aantal observaties	648	633	624	461	450	443
R ²	0.0025	0.0131	0.0167	0.0053	0.0195	0.0438
Rol opbrengstindicatoren bij beslissing zittenblijven						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Afstand tot ondergrens	0.00 (0.007)	0.01 (0.007)	0.01 (0.008)	0.00 (0.012)	0.01 (0.012)	0.02 (0.012)
Gebruik indicatoren in algemeen goede zaak		-0.03 (0.039)	-0.02 (0.043)		-0.18*** (0.061)	-0.078 (0.067)
Gebruikte indicatoren zijn de juiste			-0.01 (0.033)			-0.19*** (0.052)
Constante	1.42 (0.022)	1.48 (0.115)	1.51 (0.125)	1.98 (0.035)	2.47 (0.181)	2.65 (0.194)
Aantal observaties	1223	1190	1142	698	676	656
R ²	0.0000	0.0007	0.0009	0.0001	0.0130	0.0302

Rol opbrengstindicatoren bij doorverwijzing naar speciaal onderwijs						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.02** (0.007)	-0.01 (0.007)	-0.01 (0.007)	-0.03** (0.012)	-0.03 (0.012)	-0.01 (0.013)
Gebruik indicatoren in algemeen goede zaak		-0.06* (0.034)	-0.04 (0.038)		-0.24*** (0.058)	-0.16** (0.064)
Gebruikte indicatoren zijn de juiste			-0.03 (0.031)			-0.18*** (0.054)
Constante	1.44 (0.021)	1.60 (0.100)	1.60 (0.108)	2.088 (0.036)	2.75 (0.170)	2.97 (0.184)
Aantal observaties	1628	1577	1506	774	749	726
R ²	0.0029	0.0036	0.0033	0.0056	0.0249	0.0394

De tabel toont de coëfficiënten van een regressiemodel, met de standaardfouten tussen haakjes. Het aantal observaties in de regressie over andere scholen is lager omdat enkelen geen antwoord geven op de vraag of andere scholen vaker oefenen. Om te zien of de resultaten vergelijkbaar zijn, zijn de regressies voor de eigen school ook nog gedraaid voor alleen degenen die antwoord geven over de andere scholen. De coëfficiënten worden dan iets kleiner maar dezelfde variabelen blijven significant.

4.1.4 *Schoolleiders en leerkrachten over de rol van de opbrengstindicatoren bij de keuze voor lesmethode en leerlingvolgsysteem*

Aan schoolleiders en leerkrachten is gevraagd in hoeverre bij de aanschaf van lesmethoden gelet wordt op aansluiting hiervan bij het leerlingvolgsysteem of de eindtoets. Ongeveer een vijfde van de respondenten zegt dat het toezichtkader een heel sterke rol speelt bij deze beslissing en 40 procent een vrij sterke (zie tabel 4.12). Als gevraagd wordt of de opbrengstcriteria een rol spelen bij de aanschaf van lesmethoden op andere scholen, antwoordt 18 procent dat dit heel sterk gebeurt en 30 procent dat dit vrij sterk gebeurt.

“Met de aanschaf van lesmethodes wordt geen rekening gehouden met de eindtoets.”

“De lesmethodes moeten aansluiten bij de vraagstelling van de eindtoets, anders is het net alsof je leert autorijden zonder auto.”

“Alles staat in dienst van de opbrengsten, dus heel belangrijk dat het aansluit op de toetsen. Wij bereiden ze voor op een toets. Ook al is het niet zo bedoeld.”

Tabel 4.12 Rol opbrengstindicatoren bij aanschaf lesmethoden

In hoeverre speelt bij de keuze voor de aanschaf van lesmethoden een rol dat deze aansluiten op de eindtoets en de toetsen uit het leerlingvolgsysteem?	Op de eigen school	Op andere scholen
Helemaal niet	11.17	1.87
Een beetje	23.84	9.55
Vrij sterk	39.48	29.12
Heel sterk	21.77	17.87
Geen voortgangstoetsen	0.12	n.v.t.
Weet ik niet	n.v.t.	37.73
Antwoord ontbreekt	3.61	3.6

De tabel laat percentages van het totaal aantal respondenten (2.462) zien.

Aan leerkrachten in groep 3 tot en met groep 7 is gevraagd in hoeverre ze in de les meer aandacht besteden aan stof die op een geplande voortgangstoets aan de orde komt, als dit nog onvoldoende aanbod is geweest in de lesmethode. Tabel 4.13 laat zien dat 46 procent dit een beetje doet, en 35 procent vrij of heel sterk. Ook denkt men dat dit op andere scholen iets sterker gebeurt dan op de eigen school.

Tabel 4.13 Rol opbrengstindicatoren bij toetsen in leerlingvolgsysteem

Als u weet dat een bepaald onderdeel van een geplande toets uit het leerlingvolgsysteem op dat moment in de lesmethode nog niet voldoende aan bod is geweest, in hoeverre besteedt u daar dan extra aandacht aan in de les?		
	Op de eigen school	Op andere scholen
Helemaal niet	14.25	0.78
Een beetje	46.54	17.27
Vrij sterk	29.07	33.37
Heel sterk	6.54	18.34
Weet ik niet		26.54
Antwoord ontbreekt	3.61	3.71

De tabel laat percentages van het totaal aantal respondenten (1.025, alleen leerkrachten groep 3-7) zien.

Tabel 4.14 laat zien dat scholen die dicht bij de ondergrens zitten of die vinden dat de gehanteerde criteria de juiste zijn vaker aangeven dat op andere scholen extra aandacht aan de stof voor de voortgangstoetsen te besteden als deze aankomende zijn, of de lesmethode aanpassen aan het leerlingvolgsysteem, maar niet op de eigen school. Er is geen systematische relatie tussen de antwoorden en andere schoolkenmerken.

Tabel 4.14 Regressieresultaten: aanschaf lesmethode en relatie leerlingvolgsysteem en eindtoets.

Bij aanschaf lesmethode wordt rekening gehouden met leerlingvolgsysteem en eindtoets						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.01 (0.008)	-0.01 (0.009)	-0.01 (0.009)	-0.02* (0.009)	-0.02* (0.009)	-0.01 (0.009)
Gebruik indicatoren in algemeen goede zaak		-0.018 (0.042)	-0.00 (0.047)		-0.09* (0.045)	-0.01 (0.049)
Gebruikte indicatoren zijn de juiste			-0.06 (0.038)			-0.18*** (0.040)
Constante	2.79 (0.025)	2.84 (0.124)	2.93 (0.134)	3.100 (0.026)	3.355 (0.131)	3.57 (0.142)
Aantal observaties	1762	1704	1619	1066	1030	983
R ²	0.0014	0.0009	0.0024	0.0037	0.0066	0.0268
In les wordt extra aandacht besteed aan aankomende toets in leerlingvolgsysteem						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.02* (0.011)	-0.02 (0.012)	-0.02 (0.012)	-0.03** (0.012)	-0.03** (0.012)	-0.02* (0.012)
Gebruik indicatoren in algemeen goede zaak		0.03 (0.0592)	0.01 (0.065)		-0.10 (0.062)	0.03 (0.067)
Gebruikte indicatoren zijn de juiste			0.06 (0.050)			-0.28*** (0.053)
Constante	2.30 (0.034)	2.22 (0.172)	2.14 (0.187)	3.01 (0.036)	3.29 (0.181)	3.60 (0.198)
Aantal observaties	733	708	663	532	517	486
R ²	0.0042	0.0035	0.0053	0.0113	0.0149	0.0682

De tabel toont de coëfficiënten van een regressiemodel, met de standaardfouten tussen haakjes. Het aantal observaties in de regressie over andere scholen is lager omdat enkelen geen antwoord geven op de vraag of andere scholen vaker oefenen. Om te zien of de resultaten vergelijkbaar zijn, zijn de regressies voor de eigen school ook nog gedraaid voor alleen degenen die antwoord geven over de andere scholen. De coëfficiënten worden dan iets kleiner maar dezelfde variabelen blijven significant.

4.1.5 Schoolleiders en leerkrachten over de rol van de opbrengstindicatoren bij afname van de Cito eindtoets

Aan schoolleiders en leerkrachten van groep 8 is gevraagd hoe vaak er op school een oefentoets wordt gemaakt ter voorbereiding op de Cito eindtoets. Ongeveer de helft van de respondenten geeft aan dat er 1 of 2 keer geoefend wordt, en 15 procent geeft aan dat 3-5 keer geoefend wordt (tabel 4.15). Veel respondenten geven aan dat er in ieder geval kennis wordt gemaakt met de toets, en dat de leerlingen bekend gemaakt worden met de vraagstelling. Ook geeft men vaak aan dat er altijd met slechts delen van een toets wordt geoefend. Opvallend is dat het merendeel nu geen antwoord geeft over andere scholen. Er is gekeken of deze groep afwijkt ten opzichte van de groep die wel antwoord geeft over de andere scholen met betrekking tot de antwoorden over de eigen school maar dit is niet het geval. Degenen die wel antwoord geven over de andere scholen denken dat daar vaker geoefend wordt als op de eigen school.

“Gemiddeld wordt er eenmaal per week geoefend in de periode september - februari, dat zijn onderdelen en niet een volledige toets. In totaal zou het om twee volledige toetsen zijn van voorgaande jaren. Echter, er wordt niet geoefend om te 'scoren' maar om de kinderen te wennen aan de vraagstelling.”

“Voorgaande jaren was het een enkele keer, en eigenlijk was het geen issue. Maar met de nieuwe leerkracht hebben we besloten dat we dat meer gaan doen omdat we anders onder de norm dreigen te zakken.”

Tabel 4.15 Oefenen Cito eindtoets

Hoe vaak, schat u, maken leerlingen in groep 8 op uw school een oefentoets ter voorbereiding van de eindtoets? Kunt u het aantal keer aangeven?		
	Op de eigen school	Op andere scholen
0 keer	14.10	0.49
1 keer	31.97	5.32
2 keer	22.26	8.04
3-5 keer	14.78	10.70
6-10 keer	3.71	2.53
Meer dan 10 keer	1.74	1.92
Antwoord ontbreekt	11.44	71.00

De tabel laat percentages van het totaal aantal respondenten (1.617, alleen schoolleiders en leerkrachten groep 8) zien.

Daarnaast is aan leerkrachten van groep 8 gevraagd of ze vooraf lesstof gaan bespreken die aan de orde komt in de eindtoets. Tabel 4.16 laat zien dat een kwart van de leerkrachten zegt dit regelmatig of vaak te doen, en bijna 40 procent zegt dit nooit of bijna nooit te doen. Ongeveer 40 procent zegt dat het regelmatig of vaak op andere scholen gebeurt. Deze vraag lijkt door schoolleiders geïnterpreteerd als “we bespreken in algemene zin de lesstof waarvan we verwachten dat deze aan de orde gaat komen op de eindtoets”. We denken niet dat al deze leerkrachten de toetsboekjes van te voren in zien.

“Er wordt niet van te voren gescand wat de vragen zijn want de stof zit al in de methodieken. We zien de eindtoets ook als een bevestiging van datgene wat de school zelf doet.” “Het landelijke gemiddelde gaat omhoog dus zeer zeker dat het gebeurt. De onderwijsinspectie werkt dat zelf in de hand. Dit antwoord resonanceert in vrijwel alle antwoorden betreffende de volgende vragen over scholen in de omgeving”.

Tabel 4.16 Vooraf lesstof bespreken die aan de orde komt op de eindtoets

Hoe vaak, schat u, wordt lesstof vooraf besproken die aan de orde gaat komen in de eindtoets?		
	Op de eigen school	Op andere scholen
Nooit	26.08	3.23
Bijna nooit	14.47	7.06
Soms	28.71	26.08
Regelmatig	19.14	27.87
Vaak	6.34	12.08
Antwoord ontbreekt	5.26	23.68

De tabel laat percentages van het totaal aantal respondenten (836, alleen leerkrachten groep 8) zien.

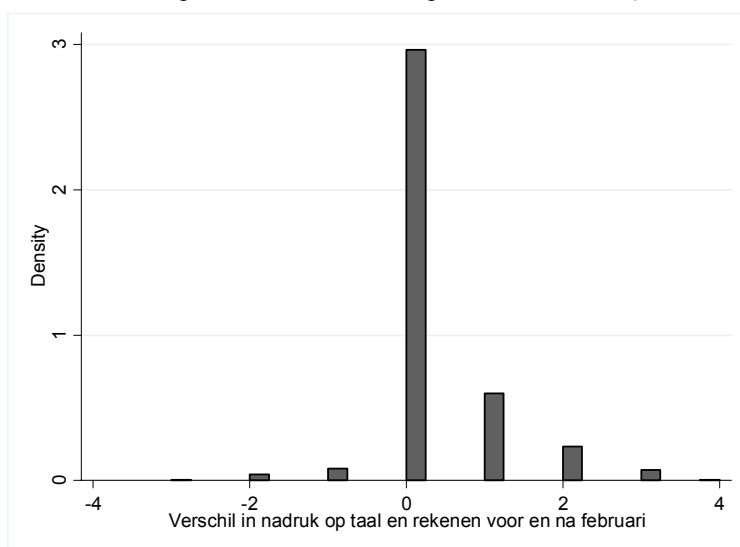
Ook is meer algemeen aan leerkrachten van groep 8 gevraagd of er in de periode voor de Cito eindtoets meer nadruk ligt op rekenen en taal in vergelijking met de periode erna. Dat wordt vaak genoemd en was ook een van de redenen om de eindtoets later in het jaar te plannen. Uit de enquête komt echter naar voren dat dit effect niet groot is. Tabel 4.17 laat zien dat er in de periode tot februari gemiddeld iets meer nadruk ligt op taal en rekenen dan in de periode na februari. Figuur 4.1 laat het verschil tussen het antwoord op de vraag tot februari minus dat op de vraag na februari zien. Positief betekent meer nadruk voor de toets.

Tabel 4.17 Nadruk op taal en rekenen voor en na de Cito eindtoets

In groep 8 is in de periode veel nadruk gelegd op taal en rekenen		
	In de periode tot februari	In de periode na februari
Zeker niet	3.47	3.83
Niet	16.39	20.81
Twijfel	5.86	11.96
Wel	46.77	44.86
Zeker wel	22.85	12.68
Antwoord ontbreekt	4.67	5.86

De tabel laat percentages van het totaal aantal respondenten (836, alleen leerkrachten groep 8) zien.

Figuur 4.1 Verschil in nadruk op taal en rekenen in de periode voor en na de Cito eindtoets (positief is meer nadruk op taal en rekenen in de periode tot februari)



Vervolgens is aan schoolleiders en leerkrachten van groep 8 gevraagd of de beoordeling van de inspectie een rol speelt bij de beslissing om leerlingen al dan niet te laten deelnemen aan de Cito eindtoets. De helft zegt dat dit helemaal niet speelt op de eigen school, maar 20 procent geeft aan dat dit toch vrij of heel sterk gebeurt (tabel 4.18). Veel respondenten merken op dat, “conform de regels”, sommige kinderen buiten de toets worden gehouden. Op andere scholen denkt men dat dit vaker gebeurt, 34 procent vrij of heel sterk.

In principe doet iedereen mee. Als het niet te belastend is voor het (zwakke) kind dat doet hij (of zij) gewoon mee want je wilt toch een beeld van hem hebben. Echter, kinderen met een rugzak bijvoorbeeld mogen van de inspectie erbuiten gelaten worden en dat gebeurt dan ook.” “Er wordt een beetje mee gerommeld.”

Tabel 4.18 Rol opbrengstindicatoren bij selectie leerlingen Cito eindtoets

In hoeverre speelt bij de keuze om een leerling wel of niet mee te laten doen aan de eindtoets de beoordeling door de inspectie een rol?	Op de eigen school	Op andere scholen
Helemaal niet	51.39	8.72
Een beetje	19.91	31.97
Vrij sterk	12.31	25.54
Heel sterk	8.04	9.96
Niet van toepassing	5.81	
Antwoord ontbreekt	2.54	23.81

De tabel laat percentages van het totaal aantal respondenten (1.617, alleen schoolleiders en leerkrachten groep 8) zien.

Vervolgens is een reeks stellingen voorgelegd aan leerkrachten van groep 8 over condities en acties tijdens afname van de Cito eindtoets. Tabel 4.19 laat zien dat ongeveer 60 procent van de groep 8-leerkrachten aangeeft dat de leerlingen gespannen waren voor de eindtoets. Enkele respondenten merken op dat dit afhangt in hoeverre er ook media aandacht aan wordt gegeven. Op de vraag of leerlingen die moeite hebben met een vraag wel eens op weg worden geholpen zegt meer dan een vijfde dat dit soms tot vaak gebeurt. Meer op de andere scholen dan op de eigen school.

“Ieder jaar is er een enorm mediahype en sommige kinderen kunnen daar niet tegen. Anderen blijven er stoïcijns onder. Werd de Citotoets niet uitvergroot in de media dan zijn ze nagenoeg niet gespannen.”
“Elk jaar is er wel één leerling, van de pakweg 31, die gekalmeerd moet worden”

Ruim 70 procent van de leerkrachten zegt leerlingen tijdens de eindtoets gerust te stellen en ruim 50 procent zegt ook dat dit op andere scholen gebeurt. Als het gaat om extra tijd geven, zegt ongeveer een vijfde dat ze dit regelmatig of vaak doen, maar een aantal leerkrachten merkt op dit alleen te doen als het conform de regels is, zoals bijvoorbeeld bij dyslectische kinderen. Ongeveer een kwart denkt dat dit op andere scholen regelmatig of vaak voorkomt. Vragen voorlezen komt minder voor, vier procent zegt dit regelmatig of vaak te doen (bijvoorbeeld, bij dyslectische kinderen). Opvallend is dat ruim 80 procent geen antwoord geeft op de vraag of men denkt dat dit op andere scholen voorkomt. Voor 60 procent zijn dit leerkrachten die zeggen het zelf nooit te doen.

Tabel 4.19 Diverse stellingen over afname Cito eindtoets

De kinderen waren erg gespannen voor de eindtoets		
	Op de eigen school	
Zeker niet	3.47	
Niet	17.34	
Twijfel	13.64	
Wel	39.35	
Zeker wel	20.22	
Antwoord ontbreekt	5.98	
Tijdens de toets leerlingen, die moeite hebben met een vraag, op weg helpen		
	Op de eigen school	Op andere scholen
Nooit	37.08	3.95
Bijna nooit	29.90	14.00
Soms	22.85	36.36
Regelmatig	4.43	17.58
Vaak	1.08	5.14
Antwoord ontbreekt	4.67	22.97
Tijdens de toets leerlingen geruststellen		
	Op de eigen school	Op andere scholen
Nooit	1.67	0.36
Bijna nooit	3.11	1.91
Soms	18.54	22.01
Regelmatig	34.81	28.35
Vaak	37.32	27.63
Antwoord ontbreekt	4.55	22.01
Tijdens de toets extra tijd geven		
	Op de eigen school	Op andere scholen
Nooit	20.45	3.95
Bijna nooit	20.22	10.29
Soms	34.21	36.96
Regelmatig	15.43	19.02
Vaak	4.78	7.06
Antwoord ontbreekt	4.90	22.73
Tijdens de toets vragen voorlezen		
	Op de eigen school	Op andere scholen
Nooit	46.77	11.60
Bijna nooit	20.45	4.07
Soms	22.97	4.43
Regelmatig	3.71	1.32
Vaak	1.08	0.36
Antwoord ontbreekt	5.02	78.23

De tabel laat percentages van het totaal aantal respondenten (836, alleen leerkrachten groep 8) zien.

Tabel 4.20 laat zien dat scholen die dichterbij de ondergrens zitten vaker zeggen: (1) te oefenen voor de eindtoets; (2) de lesstof die op de eindtoets aan de orde komt vooraf bespreken; (3) dat leerlingen gespannen zijn tijdens de eindtoets; (4) dat ze leerlingen geruststellen of op weg helpen tijdens de eindtoets; en (5) dat de criteria van de inspectie een rol spelen bij de selectie van leerlingen die deelnemen aan de eindtoets. Wat betreft het vaker oefenen voor de toets, zeggen op scholen die dichterbij de ondergrens zitten dat dit nog veel vaker voorkomt op andere scholen. De tabel laat ook

zien dat er een (matige) relatie is tussen of men het gebruik van indicatoren in het algemeen een goede zaak vindt en het oefenen voor de eindtoets, selectie leerlingen voor de eindtoets, en extra tijd geven op de eindtoets, als het gaat om andere scholen. Dit effect wordt in de meeste gevallen overgenomen als in model 3 de variabele “de gehanteerde indicatoren zijn de juiste” wordt meegenomen.

Tabel 4.20 Regressieresultaten: rol van opbrengstindicatoren bij eindtoets

Aantal keren oefenen voor eindtoets						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.07** (0.036)	-0.07** (0.037)	-0.07* (0.038)	-0.32*** (0.103)	-0.33*** (0.103)	-0.34*** (0.108)
Gebruik indicatoren in algemeen goede zaak		-0.32* (0.177)	-0.30 (0.193)		-1.51*** (0.472)	-1.09** (0.512)
Gebruikte indicatoren zijn de juiste			-0.045 (0.160)			-0.73 (0.452)
Constante	2.44 (0.110)	3.40 (0.528)	3.42 (0.560)	4.63 (0.307)	9.01 (1.399)	9.63 (1.521)
Aantal observaties	1093	1056	1016	364	351	339
R ²	0.0036	0.0073	0.0074	0.0268	0.0544	0.0630
Vooraf lesstof bespreken die aan de orde gaat komen in de eindtoets						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.05** (0.019)	-0.05** (0.020)	-0.06*** (0.021)	-0.05*** (0.017)	-0.04*** (0.017)	-0.05*** (0.018)
Gebruik indicatoren in algemeen goede zaak		-0.07 (0.093)	-0.06 (0.103)		-0.11 (0.079)	-0.07 (0.088)
Gebruikte indicatoren zijn de juiste			-0.01 (0.092)			-0.06 (0.078)
Constante	2.71 (0.060)	2.892 (0.270)	2.90 (0.291)	3.60 (0.051)	3.92 (0.231)	3.95 (0.245)
Aantal observaties	597	570	537	477	456	431
R ²	0.0092	0.0105	0.0142	0.0172	0.0191	0.0234
Verschil in nadruk op taal en rekenen in periode voor en na de eindtoets.						
	Op eigen school					
	Model 1	Model 2	Model 1			
Afstand tot ondergrens	-0.01 (0.011)	-0.01 (0.011)	-0.01 (0.012)			
Gebruik indicatoren in algemeen goede zaak		0.03 (0.050)	0.04 (0.056)			
Gebruikte indicatoren zijn de juiste			-0.00 (0.051)			
Constante	0.272 (0.034)	0.17 (0.146)	0.17 (0.158)			
Aantal observaties	591	566	534			
R ²	0.0008	0.0010	0.0014			

Rol criteria bij selectie leerlingen voor deelname aan eindtoets						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.05*** (0.011)	-0.04*** (0.011)	-0.04*** (0.011)	-0.03** (0.010)	-0.02** (0.010)	-0.01 (0.011)
Gebruik indicatoren in algemeen goede zaak		-0.165*** (0.054)	-0.06 (0.057)		-0.22*** (0.051)	-0.11* (0.056)
Gebruikte indicatoren zijn de juiste			-0.22*** (0.048)			-0.26*** (0.047)
Constante	1.82 (0.034)	2.29 (0.160)	2.51 (0.168)	2.53 (0.032)	3.16 (0.152)	3.43 (0.160)
Aantal observaties	1134	1097	1056	936	904	868
R ²	0.0171	0.0235	0.0402	0.0069	0.0254	0.0561
Leerlingen waren gespannen voor de toets						
	Op eigen school					
	Model 1	Model 2	Model 3			
Afstand tot ondergrens	-0.07*** (0.017)	-0.07*** (0.017)	-0.07*** (0.018)			
Gebruik indicatoren in algemeen goede zaak		-0.10 (0.080)	-0.03 (0.089)			
Gebruikte indicatoren zijn de juiste			-0.16** (0.080)			
Constante	3.67 (0.052)	3.93 (0.231)	4.08 (0.251)			
Aantal observaties	598	571	538			
R ²	0.0304	0.0314	0.0387			
Tijdens de eindtoets leerlingen die moeite hebben met een vraag op weg helpen						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.03** (0.015)	-0.03** (0.015)	-0.04*** (0.016)	-0.03** (0.015)	-0.03* (0.016)	-0.03 (0.017)
Gebruik indicatoren in algemeen goede zaak		-0.01 (0.069)	-0.06 (0.076)		-0.05 (0.075)	0.01 (0.084)
Gebruikte indicatoren zijn de juiste			0.11 (0.069)			-0.085 (0.073)
Constante	2.00 (0.045)	2.02 (0.201)	1.93 (0.214)	3.14 (0.047)	3.30 (0.217)	3.32 (0.230)
Aantal observaties	602	575	541	479	459	434
R ²	0.0091	0.0084	0.144	0.0089	0.0090	0.0111
Tijdens de eindtoets leerling geruststellen						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.05*** (0.014)	-0.04*** (0.014)	-0.05*** (0.015)	-0.04*** (0.014)	-0.04*** (0.014)	-0.04*** (0.015)
Gebruik indicatoren in algemeen goede zaak		-0.02 (0.064)	0.05 (0.071)		-0.09 (0.067)	-0.03 (0.076)
Gebruikte indicatoren zijn de juiste			-0.11* (0.064)			-0.08 (0.066)
Constante	4.15 (0.043)	4.19 (0.187)	4.28 (0.201)	4.09 (0.043)	4.35 (0.195)	4.37 (0.208)
Aantal observaties	602	575	541	486	465	439

R ²	0.0214	0.0174	0.0283	0.0198	0.0228	0.0279
Leerlingen tijdens de eindtoets extra tijd geven						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.01 (0.017)	-0.00 (0.017)	-0.01 (0.018)	-0.00 (0.015)	0.00 (0.016)	0.00 (0.017)
Gebruik indicatoren in algemeen goede zaak		-0.17** (0.080)	-0.12 (0.088)		-0.19*** (0.075)	-0.09 (0.084)
Gebruikte indicatoren zijn de juiste			-0.06 (0.079)			-0.15** (0.073)
Constante	2.67 (0.053)	3.15 (0.231)	3.16 (0.248)	3.24 (0.047)	3.79 (0.217)	3.86 (0.229)
Aantal observaties	599	572	539	482	461	435
R ²	0.0004	0.0080	0.0080	0.0000	0.0143	0.0206
Leerlingen tijdens de eindtoets vragen voorlezen						
	Op eigen school			Op andere scholen		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Afstand tot ondergrens	-0.01 (0.015)	-0.01 (0.016)	-0.01 (0.016)	-0.03 (0.031)	-0.03 (0.032)	-0.03 (0.033)
Gebruik indicatoren in algemeen goede zaak		-0.12* (0.071)	-0.13* (0.079)		-0.23 (0.161)	-0.10 (0.173)
Gebruikte indicatoren zijn de juiste			-0.03 (0.071)			-0.32** (0.151)
Constante	1.86 (0.047)	2.22 (0.207)	2.29 (0.222)	1.87 (0.099)	2.55 (0.471)	2.89 (0.490)
Aantal observaties	600	573	540	137	130	122
R ²	0.0010	0.0058	0.0084	0.0075	0.0249	0.0634

De tabel toont de coëfficiënten van een regressie model, met de standaardfouten tussen haakjes. Het aantal observaties in de regressie over andere scholen is een stuk lager, omdat velen geen antwoord geven op de vraag of andere scholen vaker oefenen. Om te zien of de resultaten vergelijkbaar zijn, zijn de regressies voor de eigen school ook nog gedraaid voor alleen degenen die antwoord geven over de andere scholen. De coëfficiënt van de variabele afstand tot de ondergrens is dan ongeveer half zo groot als die voor andere scholen en ook significant.

4.1.6 Verschillen tussen schoolleiders en leerkrachten

Daar waar vragen aan zowel schoolleiders als leerkrachten zijn gesteld, is ook gekeken naar verschillen tussen beide groepen. Leerkrachten geven aan minder goed op de hoogte te zijn van het toezichtkader van de inspectie dan schoolleiders: 94 procent van de schoolleiders zegt het toezichtkader goed of zeer goed te kennen tegenover 71 procent van de leerkrachten. Ook twijfelen leerkrachten vaker of de gehanteerde indicatoren de juiste zijn: zo vindt 58 procent van de schoolleiders de indicatoren de juiste, tegenover 36 procent van de leerkrachten. Ongeveer een vijfde van de leerkrachten vindt ook dat het gebruik van opbrengstindicatoren een negatief effect heeft op het onderwijs op school, tegenover minder dan tien procent van de ondervraagde schoolleiders.

De vragen over de rol van de opbrengstindicatoren bij zittenblijven, doorverwijzing naar het speciaal onderwijs of aanschaf van de lesmethode zijn ook aan beide groepen gesteld. Leerkrachten geven vaker aan dat de opbrengstindicatoren een vrij sterke of zeer sterke rol spelen bij deze beslissingen, en zeggen ook vaker dat het op andere scholen voorkomt. De overige vragen zijn aan leerkrachten en schoolleiders die zelf lesgeven gesteld, maar de aantallen bij de laatste groep zijn te klein om daar robuuste uitspraken over te doen.

4.1.7 De resultaten van schoolleiders en leerkrachten samengevat

Alles bij elkaar komt het volgende beeld voren. Bijna alle vormen van mogelijk strategisch gedrag die bevraagd zijn, komen volgens de leerkrachten en schoolleiders voor. De laagste percentages zitten rond de 20 procent voor het niet toelaten van zwakke leerlingen, zittenblijven en doorverwijzen naar speciaal onderwijs. Voor al deze thema's geeft men overigens aan dat dit op andere scholen veel vaker voorkomt. Dat zou er op kunnen wijzen dat men het moeilijk vindt om toe te geven dat deze vormen van strategisch gedrag voorkomen.

Bij de volgende aspecten geven leerkrachten en schoolleiders op bijna alle scholen aan dat dit voor komt: het meerdere keren oefenen van de eindtoets, vooraf lesstof bespreken die in de eindtoets aan de orde komt, het op weg helpen van leerlingen bij de eindtoets, geruststellen van leerlingen bij de eindtoets en ze extra tijd geven om deze toets te maken. Met uitzondering van het geven van de extra tijd, komen al deze aspecten meer voor bij scholen die lager scoren met betrekking tot de norm voor opbrengsten op de eindtoets.

Over het algemeen geeft men aan dat de genoemde vormen van strategisch gedrag mogelijk meer voorkomen op andere scholen. Uitzonderingen hierop zijn het aanschaffen van de lesmethode die aansluit op toetsen, het oefenen van de eindtoets en het geruststellen van leerlingen. Kennelijk worden dit als algemeen aanvaardbaar gedrag gezien. Steeds zien we ook dat leerkrachten meer dan schoolleiders aangeven dat deze vormen van gedrag voorkomen op de school. Hierbij zou het het geval kunnen zijn dat schoolleiders niet op de hoogte zijn van dit gedrag. Voor schoolleiders die zelf ook les geven geldt echter ook dat zij minder vaak aangeven dat de gevraagde vormen van gedrag voorkomen op de school. Er lijkt dus sprake te zijn van meer sociaal wenselijke antwoorden door schoolleiders.

Voor meer dan de helft van de gevraagde vormen van gedrag vinden we een verband met de afstand die een school gemiddeld heeft tot de ondergrens voor opbrengsten die de inspectie hanteert. Dat kan erop duiden er op dat deze gedragingen mede het gevolg zijn van strategische overwegingen van de school om deze score te beïnvloeden. Vergeleken met de mate waarin deze vormen van gedrag voorkomen is het verschil tussen scholen die er belang bij hebben en scholen die er geen belang bij hebben echter klein.

De onderzochte vormen van mogelijk strategisch gedrag hangen in een aantal gevallen ook samen met de kijk die de school heeft op het gebruik van indicatoren en de specifieke keuze van deze indicatoren. Dat geldt voor de toelating van zwakke leerlingen, de aanschaf van lesmethodes die passen bij de toetsen, het vooraf bespreken van lesstof die in de toets aan de orde gaat komen, selectie van leerlingen voor de toets, en extra tijd tijdens de toets.

4.2 Ouders

4.2.1 Aanpak en respons ouders

In Zuid-Limburg worden jaarlijks gegevens verzameld op basisscholen waarbij ouders worden ondervraagd naar hun ervaringen met de basisschool en hun schoolkeuze. Deze gegevensverzameling maakt deel uit van de Onderwijsmonitor Limburg waarbij scholen, besturen en de Universiteit gezamenlijk de voortgang van leerlingen in de regio monitoren, de dialoog aangaan over

geconstateerde knelpunten en daarvoor oplossingen zoeken³⁹. Voor dit onderzoek gebruiken we gegevens uit de oudervragenlijst van groep 2 over schoolkeuze en uit groep 8 over zittenblijven, doorverwijzing en de ervaringen met de eindtoets. De informatie die beschikbaar is over deze Zuid-Limburgse ouders is meer gedetailleerd dan informatie die beschikbaar is over ouders in de rest van Nederland. Deze groep ouders is waarschijnlijk representatief voor andere Nederlandse ouders. Ervaring hiermee laat weinig verschillen zien bij gegevens die voor alle groepen beschikbaar zijn. Limburg kent overigens relatief weinig zwakke scholen.

In het databestand van groep 2 zitten 2.544 leerlingen, waarvan 2.399 ouders (94,3 procent) de vragenlijst hebben ingevuld. 6 ouders vulden het formulier in zonder naam, waardoor de gegevens niet gekoppeld konden worden en de overige ouders vulden niets in. Van de 2.399 ouders die meededen, vulde 2,9 procent in dat ze niet willen dat de gegevens van hun kind mee gaan in het onderzoek. Dit levert een uiteindelijke steekproef op van 2.325 ouders.

In het databestand van groep 8 zitten 5.163 leerlingen, waarvan 3.289 ouders (63,7 procent) de vragenlijst hebben ingevuld. De rest heeft niet geantwoord. Van deze 3.289 ouders geeft 4,0 procent aan dat ze niet willen dat de gegevens van hun kind mee gaan in het onderzoek. Dit levert een uiteindelijke steekproef op van 3.159 ouders. Om na te gaan in hoeverre de ouders en leerkrachten verschillen ten aanzien van vragen rondom de eindtoets (deze vragen zijn aan beide groepen gesteld) zijn de databestanden van Zuid-Limburg en die uit het survey voor dit onderzoek aan elkaar gekoppeld. Dit levert 749 matches op, verdeeld over 40 scholen. Dat is 23 procent van het aantal scholen in de Zuid-Limburgse data. Tegenover 749 ouders staan dan 189 schoolleiders en 560 leerkrachten.

4.2.2 Ouders over schoolkeuze/toelating

Aan de ouders van leerlingen van groep 2 zijn enkele vragen gesteld over schoolkeuze. Eerst is gevraagd of men uit een aantal scholen kon kiezen. Tabel 4.21 laat zien dat ongeveer de helft van de ouders aangeeft dat ze konden kiezen voor drie of meer scholen. 13 procent zegt dat er geen keuze was. Ruim 53 procent van de ouders geeft overigens aan niet veel informatie te hebben ingewonnen over de school voordat ze een keuze maakten, 42 procent zegt dit wel te hebben gedaan.

Ruim 91 procent van de ouders zegt dat hun kind nu op de school van hun eerste keuze zit, 6 procent geeft aan dat dit niet zo is. Redenen waarom het kind niet op de school van eerste keuze zit zijn: 18 procent school is opgeheven of gefuseerd met school die ouders niet aanstond, 18 procent afstand en bereikbaarheid met vervoer, 9 procent door verhuizing van de ouders, 28 procent voelde niet goed na een jaar of na de kijkdag. Een kwart geeft andere, zeer uiteenlopende redenen, zoals geloof, of te dorps, of geen jenaplan of Dalton school. Dit zijn dus met name heroverwegingen. Niet worden toegelaten wordt niet genoemd.

³⁹ Voor meer informatie zie www.kaans.nl.

Tabel 4.21 Ouders in groep 2 over schoolkeuze

Uit hoeveel basisscholen kon u kiezen	
Ik kon niet kiezen	13.16
Ik kon kiezen uit twee scholen	20.39
Ik kon kiezen uit drie of meer scholen	51.44
Weet ik niet	13.20
Antwoord ontbreekt	1.81
Was deze school (kind zit nu in groep 2) uw eerste keuze?	
Ja	91.14
Nee	6.11
Antwoord ontbreekt	2.75

De tabel laat percentages van het totaal aantal respondenten (2.325 ouders groep 2) zien.

4.2.3 Ouders over zittenblijven of doorverwijzing naar speciaal onderwijs

Aan ouders van leerlingen in groep 8 zijn een aantal vragen gesteld over zittenblijven. Tabel 4.22 zet de vraag “heeft de school u wel eens aangeraden dat uw kind een jaar over moest doen” uitgezet tegen de mening van de ouder dat het kind wat hem/haar betreft beter een jaar over had kunnen doen. In 84 procent van de gevallen dat de ouder niet vindt dat het kind beter een jaar over had kunnen doen, heeft de school dit ook niet aangeraden. In 10 procent van de gevallen waar de ouder niet vindt dat het kind beter een jaar over had kunnen doen, heeft de school dit toch aangeraden. Van deze leerlingen (266 in totaal) is uiteindelijk 76 procent blijven zitten, dan wel in groep 2 (130 leerlingen) of hoger (73 leerlingen).

In ruim 56 procent van de gevallen waarbij de ouder vindt dat het kind wat hem/haar betreft beter een jaar had kunnen overdoen, heeft de school toch aangeraden dit niet te doen. Van deze leerlingen (122 in totaal) is uiteindelijk 9 procent blijven zitten, dan wel in groep 2 (3 leerlingen) of hoger (8 leerlingen). Het oordeel van de school bij zittenblijven weegt kennelijk zwaarder dan dat van de ouder.

Tabel 4.22 De school raadde de ouders aan dat hun kind een jaar over moest doen (kolom percentages)

	Wat ons betreft had ons kind beter een jaar langer over de basisschool kunnen doen				
	Nee	Twijfel	Ja	Geen antwoord	Aantal respondenten
Nee	84.40	74.38	56.48	10.29	2338
Twijfel	0.94	5.63	1.85	1.18	40
Ja	10.89	17.50	37.96	20.00	444
Geen antwoord	3.77	2.50	3.70	68.53	337
Aantal respondenten	2443	160	216	340	3159

De tabel laat percentages van het totaal aantal respondenten (3.159 ouders groep 8) zien.

Daarnaast is de ouders gevraagd of er ooit sprake is geweest dat hun kind de overstap zou maken naar het speciaal onderwijs (tabel 4.23). In ongeveer negen procent van de gevallen waarin de ouder aangeeft dat hun kind een zorgleerling is, is dit door de school ook geopperd. In meer dan 70 procent van de gevallen waarin de ouder aangeeft dat hun kind een zorgleerling is, is dit echter niet geopperd.

Tabel 4.23 Ons kind zou tijdens de basisschool mogelijk de overstap maken naar speciaal onderwijs (kolom percentages)

	Ons kind is een zorgleerling vanwege leer- of gedragsmoeilijkheden				Aantal respondenten
	Ja	Nee	Weet niet	Geen antwoord	
Nee	72.60	84.69	74.32	55.32	2601
Twijfel	11.19	0.92	10.81	4.26	83
Ja	8.68	0.62	1.35	0.00	55
Geen antwoord	7.53	13.77	13.51	40.43	420
Aantal respondenten	438	2600	74	47	3159

De tabel laat percentages van het totaal aantal respondenten (3.159 ouders groep 8) zien.

4.2.4 Ouders over de Cito eindtoets

Aan de ouders van leerlingen van groep 8 is ook gevraagd of hun zoon/dochter heeft deelgenomen aan de Cito eindtoets: 94,9 procent zegt van wel, 2,8 procent zegt van niet en 2,3 procent van de ouders weet het niet. Als de leerling niet heeft deelgenomen aan de Cito eindtoets, is de ouders gevraagd naar de reden hiervoor. In 40 procent van de gevallen hanteerde de school een andere toets dan de Cito eindtoets (in de helft van deze gevallen de drempeltoets, in de andere helft andere toetsen waarvan de ouders de naam niet weten, zo wordt wel DMT genoemd). In een kwart van de gevallen zegt de ouder dat de Cito toets nog moet plaatsvinden omdat de school de digitale versie afneemt. Deze is later. In zeven procent van de gevallen deed de leerling niet mee omdat er sprake was van (ernstige) dyscalculie of dyslexie.

Vervolgens zijn een aantal stellingen aan de ouders voorgelegd over de Cito eindtoets. Tabel 4.24 laat de resultaten zien. Ruim 80 procent van de ouders geeft aan dat er op school geoefend is voor de Cito eindtoets. Als we de scholen bekijken waarvoor we zowel de antwoorden van ouders als die van leerkrachten hebben, zien we dat 86 procent van de ouders zegt dat er veel geoefend is, tegenover 51 procent van de leerkrachten die zegt dat er meer dan één keer geoefend is. Ouders hebben de indruk dat er vaker geoefend wordt. Ongeveer 38 procent van de ouders geeft aan dat hun kind erg gespannen was voor de Cito eindtoets. Op de scholen waarvoor we deze informatie hebben voor zowel ouders als leerkrachten antwoordt 43 van zowel de ouders als de leerkrachten dat de leerlingen gespannen waren. Hierover is men het eens. Ruim 80 procent van de ouders geeft wel aan dat de sfeer op school tijdens de afname van de Cito eindtoets prettig was. Tot slot geven de ouders aan dat zowel in de periode tot als na februari veel aandacht wordt besteed aan taal en rekenen. Ook bij hen, net als bij de leerkrachten van groep 8, zien we dat er iets meer nadruk ligt op taal en rekenen in de periode tot februari in vergelijking met de periode erna. Figuur 4.2 toont dit grafisch. De grafiek is compacter uit dan die bij leerkrachten groep 8 (figuur 4.1). Leerkrachten geven dus vaker dan ouders aan dat er een verschil is in nadruk voor en na de eindtoets.

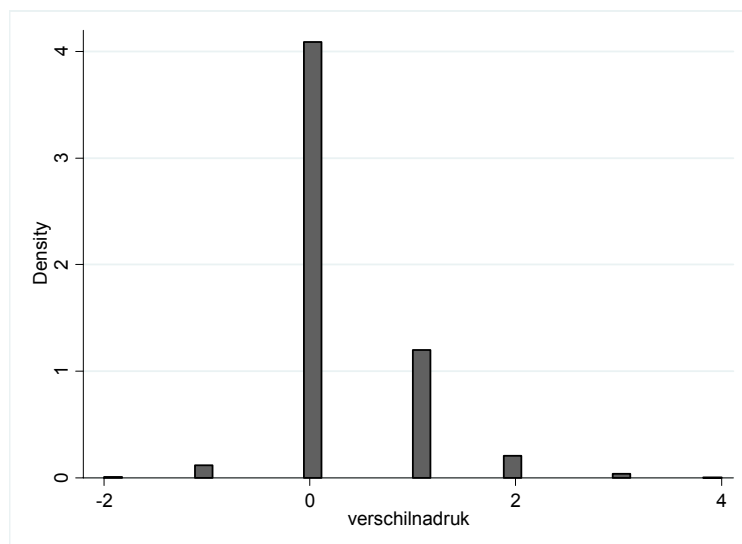
Tabel 4.24. Stellingen aan ouders rondom Cito

Op school is veel geoefend voor de Cito eindtoets	
Zeker niet	0.51
Niet	4.18
Twijfel	9.02
Wel	49.76
Zeker wel	33.90

Antwoord ontbreekt	2.63	
Mijn kind was erg gespannen voor de Cito eindtoets		
Zeker niet	8.61	
Niet	32.07	
Twijfel	16.97	
Wel	26.08	
Zeker wel	12.95	
Antwoord ontbreekt	3.32	
Tijdens het afnemen van de Cito eindtoets was de sfeer op school prettig		
Zeker niet	0.41	
Niet	1.14	
Twijfel	9.05	
Wel	64.17	
Zeker wel	18.99	
Antwoord ontbreekt	6.24	
In groep 8 lag in de periode veel nadruk op taal en rekenen		
	Tot februari	Na februari
Zeker niet	0.28	0.63
Niet	3.45	6.74
Twijfel	22.54	35.14
Wel	54.07	43.18
Zeker wel	15.23	8.48
Antwoord ontbreekt	4.43	5.82

De tabel laat percentages van het totaal aantal respondenten (3159 ouders groep 8) zien.

Figuur 4.2 Verschil in nadruk op taal en rekenen in de periode voor en na de Cito eindtoets (positief is meer nadruk op taal en rekenen in de periode tot februari)



4.2.5 Ouders over het advies voor voortgezet onderwijs

Tot slot zijn aan de ouders enkele vragen gesteld over het advies dat de leerling kreeg voor het voortgezet onderwijs en hoe tevreden men daarmee was. Tabel 4.25 laat zien dat de meeste ouders tevreden zijn met het advies dat hun kind kreeg voor het vervolgonderwijs. Ook zien we dat hoe hoger het advies, hoe meer ouders tevreden zijn. Ouders hebben over het algemeen ook het gevoel dat ze

inspraak hebben bij de totstandkoming van het advies: 80 procent geeft aan dat gevoel te hebben, 3 procent geeft aan dit gevoel niet te hebben. Deze laatste groep bestaat uit ouders van leerlingen met alle adviezen, maar iets vaker betreft het ouders van leerlingen met een vmbo theoretisch advies (20 procent) of havo advies (20 procent).

Tabel 4.25 Tevredenheid van ouders met advies voor voortgezet onderwijs (uitgesplitst naar advies, rij percentages)

Advies	% dat dit advies kreeg	Ontevreden	Geen oordeel	Tevreden	Geen antwoord	Aantal respondenten
Vmbo basis-kader	7.66	13.64	15.29	68.60	2.48	242
Vmbo bk / theoretisch	8.93	11.35	15.96	70.92	1.77	282
Vmbo theoretisch	15.51	8.57	9.18	80.61	1.63	490
Vmbo theor. / havo	6.08	9.38	9.90	78.65	2.08	192
Havo	13.83	4.58	3.43	91.30	0.69	437
Havo / vwo	17.25	5.87	2.94	89.54	1.65	545
Vwo	21.40	4.59	0.44	94.23	0.74	676
Geen antwoord	9.34	6.78	29.83	37.63	25.76	295
Aantal respondenten	3159	228	268	2547	116	3159

De tabel laat percentages van het totaal aantal respondenten (3159 ouders groep 8) zien.

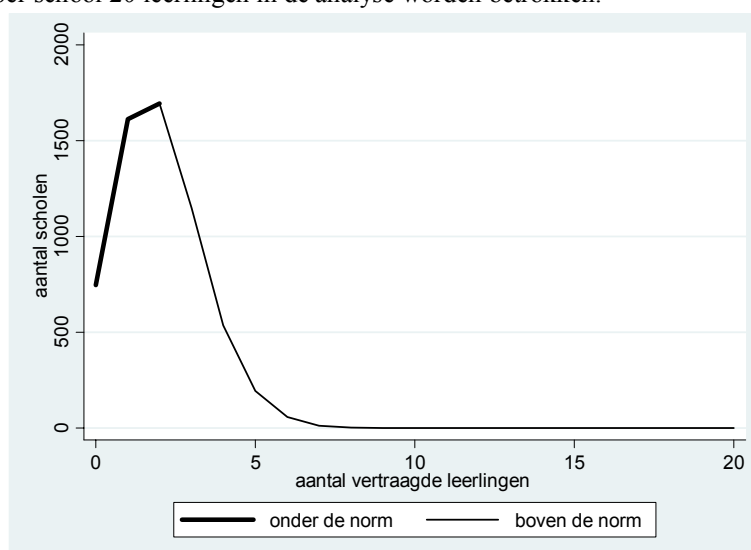
5 Analyse van twee opbrengstindicatoren op basis van microdata

5.1 Vertraagde doorstroom

Stel een school ziet goed aan de ontwikkeling van leerlingen welke leerling moet blijven zitten en welke niet. Volgens het toezichtkader van de inspectie zou aan het einde van de basisschool 15 procent van de leerlingen studievertraging op mogen lopen. 12 procent vanwege kleutergroep-verlenging in groep 3 en 3 procent voor zittenblijven in groep 3-8. In deze paragraaf bekijken we in hoeverre strategisch gedrag dat het gevolg kan zijn van een dergelijke indicator, tot uiting zou kunnen komen in grootschalige databestanden, dat zijn bestanden waarin enkele gegevens over de schoolloopbaan van alle leerlingen zijn opgenomen. Alvorens deze data te analyseren bekijken we eerst theoretisch welke effecten verwacht kunnen worden door een simulatie te maken van een hypothetische groep scholen. Bij wijze van voorbeeld gaan we er van uit dat het gemiddeld 10 procent van de leerlingen vanuit onderwijsperspectief gunstig is om vertraging op te lopen gedurende de acht jaar. Gemiddeld genomen zit dit dus onder de 15%-norm van de inspectie. Als echter de kans dat een leerling baat heeft bij vertraagde doorstroom 10 procent is – en deze kans gelijk is voor alle scholen – dan zal in sommige gevallen door het toeval, het percentage leerlingen dat vertraging oploopt, boven de 15 procent uitkomen. In sommige andere gevallen zal dit percentage door het toeval veel lager zijn dan 10 procent. De norm kan daarmee restrictief worden voor de school. Als de school onder druk van de norm er voor zorgt dat het percentage vertraagde leerlingen onder de 15 procent blijft, zullen soms leerlingen over gaan naar de volgende klas, terwijl het voor hen beter was geweest als dat niet was gebeurd. Een percentage van 20 procent vertraagde doorstroom kan dan af en toe goed gedrag zijn, dat beperkt gaat worden door die regel. De norm houdt geen rekening met toeval of spreiding rond het gemiddelde.

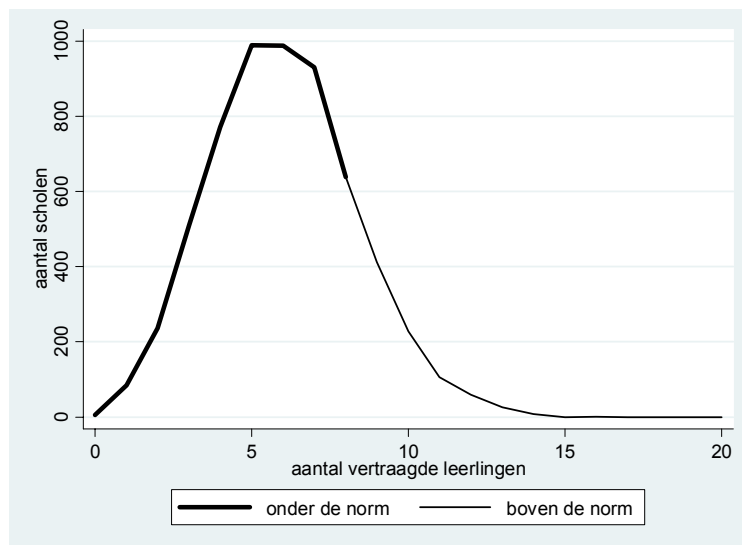
Figuur 5.1 laat zien wat de verdeling van zittenblijvers zou zijn bij 6.000 basisscholen met elk 20 leerlingen in een klas, waarbij de kans per leerling op vertraging 10 procent is. 68 procent van de scholen voldoet aan de norm van 15 procent, 32 procent voldoet daar niet aan.

Figuur 5.1 Gesimuleerde verdeling van het aantal vertraagde leerlingen per school als de kans per kind 10% is en er per school 20 leerlingen in de analyse worden betrokken.



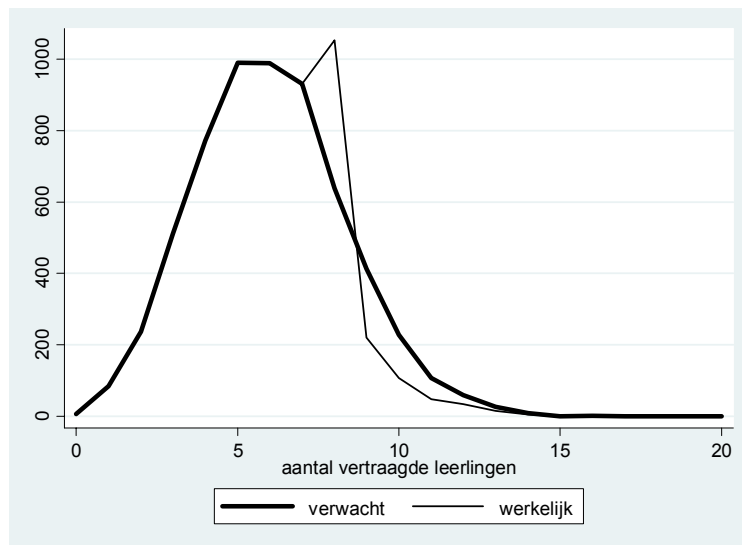
Een methode om meer nauwkeurigheid in de indicatoren te krijgen is om meerdere jaren samen te nemen. Figuur 5.2 laat zien hoe de verdeling van vertraagde leerlingen per school in het fictieve voorbeeld er uit zou zien als niet 20 maar $3 \times 20 = 60$ leerlingen per school in de indicator worden meegenomen. Met extra jaren waarnemingen, blijkt nu 86 procent van de scholen, puur op basis van toevallige variatie, binnen de norm van 15 procent te vallen.

Figuur 5.2 Gesimuleerde verdeling van het aantal vertraagde leerlingen per school als de kans per kind 10% is en er per school 60 leerlingen in de analyse worden betrokken.



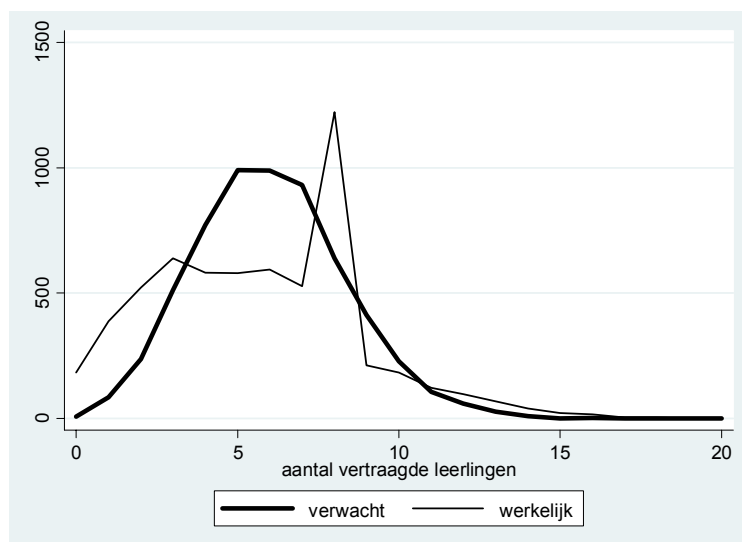
Stel nu dat het overschrijden van de norm van 15 procent negatieve consequenties kan hebben voor een school. Het kan dan zijn dat scholen zich strategisch gaan gedragen en er voor zorgen dat ze aan de norm voldoen. Figuur 5.3 laat zien wat er gebeurt als de helft van de scholen door strategisch gedrag ervoor zorgt dat ze aan de gestelde norm voldoen. Per individuele school kan dan niet worden vastgesteld of de school zich strategisch gedraagt. Net als bij het voorbeeld met de dobbelsteen in hoofdstuk 2 kan het puur toeval zijn dat een school net negen vertraagde leerlingen heeft en dus net aan de norm van 15 procent voldoet. Maar figuur 5.3 laat ook zien dat er meer scholen zijn die precies negen vertraagde leerlingen hebben, dan verwacht mocht worden op basis van het toeval. Deze toename bij negen leerlingen gaat volledig ten koste van scholen met meer dan negen vertraagde leerlingen. Dit maakt het plausibel dat de reden voor de afwijking in het patroon inderdaad is dat de school aan de norm van 15 procent probeert te voldoen. Bij de lage aantallen vertraagde leerlingen is deze hypothetische grafiek exact gelijk aan de verwachte verdeling.

Figuur 5.3 Gesimuleerde verdeling van het aantal vertraagde leerlingen per school als de helft van de scholen strategisch gedrag vertoont.



Een complicatie is dat er in de praktijk verschillen zullen zijn tussen scholen in de gemiddelde kans dat een leerling blijft zitten. Figuur 5.4 laat de consequentie zien. De piek bij negen leerlingen is nog steeds te zien, maar nu wijkt de grafiek ook bij lagere leerlingaantallen af van de verwachting. De reden hiervoor is dat de vergelijkingsgrafiek gebaseerd is op de veronderstelling dat ieder kind dezelfde kans op vertraging heeft, terwijl deze kans feitelijk varieert per school. Op basis van gegevens over meerdere jaren kan rekening worden gehouden met deze variatie tussen scholen. Verwacht kan overigens worden, dat niet alle scholen die aan de norm proberen te voldoen, exact op de norm zullen gaan zitten, maar dat in de piek ook meer spreiding zal zitten.

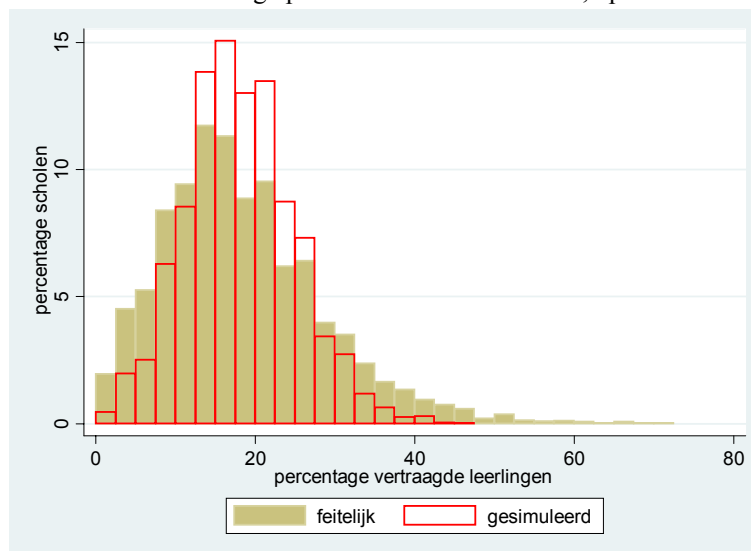
Figuur 5.4 Gesimuleerde verdeling van het aantal vertraagde leerlingen per school als er door toeval 9 vertraagde leerlingen op een school zitten in vergelijking met situatie waarin de helft van de scholen strategisch gedrag vertoont.



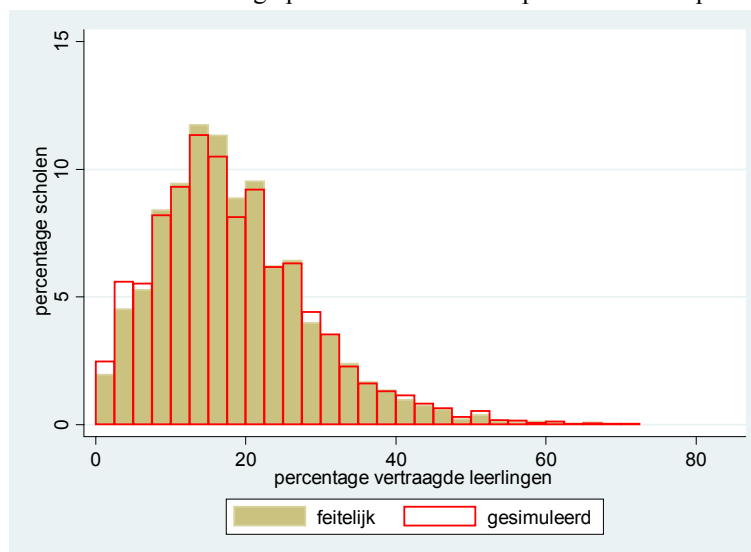
Op basis van data over vertraagde doorstroom van alle leerlingen in de schooljaren 2009/10-2012/13 kan de werkelijke verdeling van vertraagde doorstroom worden gemaakt. Hiervoor gebruiken we de registergegevens van DUO voor de jaren 2009/10-2012/13, waarin van elke leerling

die voor het laatste jaar in groep 8 van de basisschool zit, bekend is welke leeftijd hij of zij op 1 oktober van dat schooljaar had. Leerlingen die dan ouder zijn dan 11 worden gezien als vertraagde doorstroom. Hierbij kijken we alleen naar scholen met op zijn minst 20 leerlingen in groep 8. De gemiddelde kans op vertraagde doorstroom is van deze leerlingen is 18,1 procent, dat is dus hoger dan de norm van de inspectie. Figuur 5.5 geeft de verdeling van het percentage per school per jaar weer. Als we er van uit gaan dat ieder kind 18,1 procent kans heeft op vertraagde doorstroom kan gesimuleerd worden, hoe de verdeling van het percentage vertraagde leerlingen per school er uit zou zien. De figuur maakt duidelijk dat de feitelijke verdeling meer spreiding kent dan de gesimuleerde verdeling. Kennelijk bestaan er aanzienlijke verschillen in de kans op vertraagde doorstroom per school. Uit een analyse blijkt dat de helft van de variatie in het percentage vertraagde leerlingen toeval is en de helft van de variatie een schoolkenmerk is. We hebben daarom dezelfde simulatie nogmaals uitgevoerd, waarbij per school de kans op vertraagde doorstroom over de afgelopen vier jaar als uitgangspunt is genomen. Omdat dat een vrij korte periode is zal de ingeschatte kans deels afhankelijk zijn van de toevallige fluctuaties in de vertraagde doorstroom. Figuur 5.6 geeft deze verdeling.

Figuur 5.5 Feitelijke verdeling van het percentage leerlingen met vertraagde doorstroom, 2009/10-2012/13 en een simulatie van die verdeling op basis van een kans van 18,1 procent kans op vertraging.



Figuur 5.6 Feitelijke verdeling van het percentage leerlingen met vertraagde doorstroom, 2009/10-2012/13 en een simulatie van die verdeling op basis van de schoolspecifieke kans op vertraging.



Bij deze analyse is de overeenkomst tussen de gesimuleerde en de feitelijke verdeling veel groter. Een opmerkelijk verschil zit echter bij de kans dat 0-5 procent van de leerlingen en 12.5-22.5 procent van de leerlingen in een bepaald jaar op een bepaalde school vertraagd is. In de eerste categorie komt minder vaak voor dan verwacht, de tweede juist vaker. Het gaat in beide gevallen om ongeveer twee procent van de scholen. Hoewel de omvang niet heel groot is, zijn de verschillen duidelijk significant. Het kan dus geen toeval zijn. Scholen die door het toeval heel weinig vertraagde leerlingen zouden hebben, besluiten sneller leerlingen een jaar langer over groep 2 te laten doen of laten sneller leerlingen zitten. Omdat van 12,5 procent vertraagde leerlingen juist een overschot wordt geconstateerd, lijken de zone van 0-12,5 procent de zone te zijn waarin sneller besloten wordt om leerlingen langer over de basisschool te laten doen. Het is uit de grafiek niet op te maken of twee procent van de scholen zo'n tien procent meer vertraagde leerlingen heeft dan men zou verwachten, of dat bijvoorbeeld tien procent twee procent meer vertraging heeft.

5.2 De eindtoets basisonderwijs

De belangrijkste indicator in het opbrengstkader is de eindtoets basisonderwijs. De beoordeling van de resultaten die een school behaalt voor de eindtoets basisonderwijs, is vrij bepalend voor het algemene oordeel over de school. De meeste scholen gebruiken de Cito eindtoets. Afhankelijk van het percentage gewichtenleerlingen op de school heeft de inspectie grenzen vastgesteld. Als een school in drie opeenvolgende jaren deze ondergrens niet haalt, worden de eindopbrengsten als onvoldoende beoordeeld. Het valt dus te verwachten dat scholen zich zullen inspannen om deze norm te halen.

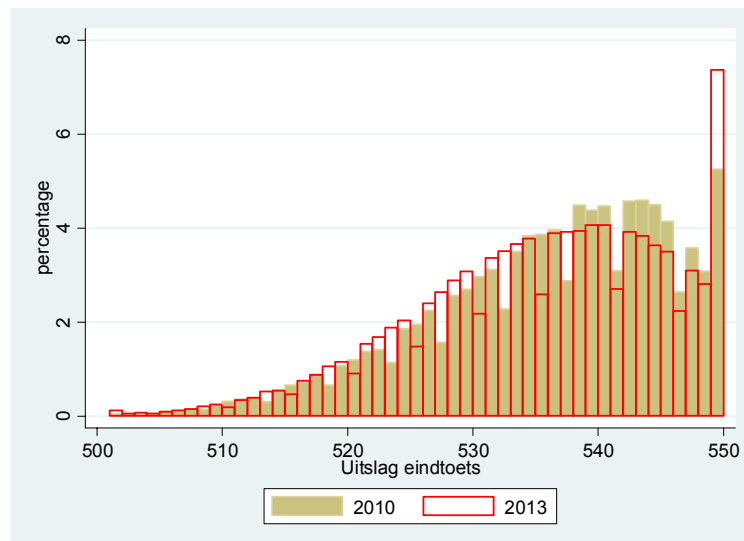
In deze paragraaf analyseren we, net als bij de indicator vertraagde doorstroom, of uit de verdeling van de scores indicaties kunnen worden ontleend over het gedrag van de scholen. Om twee redenen is dat lastiger dan de analyse van de vertraagde doorstroom:

In de eerste plaats is de eindscore van leerlingen een continue variabele en geen 0-1-variabele. Dat betekent dat de omvang van toevallige fluctuaties in deze indicator niet afgeleid kunnen worden uit de gemiddelde omvang. Die moeten dus apart in beschouwing worden genomen.

Daarnaast volgt uit de hoogte van de score op de eindtoets niet hoe deze resultaten tot stand zijn gekomen. Toetsscores kunnen hoger zijn doordat de kwaliteit van het onderwijs is verbeterd, doordat de school de focus heeft gelegd op de lesinhoud die getoetst wordt, of doordat de school maatregelen neemt die de toetsscore zelf gunstig beïnvloeden. In de analyse in deze paragraaf kunnen we alleen bekijken of de toetsscores beïnvloed lijken te worden door het toezichtkader van de inspectie. In een volgend hoofdstuk gaan we in op de het gedrag van scholen dat hierop van invloed kan zijn geweest.

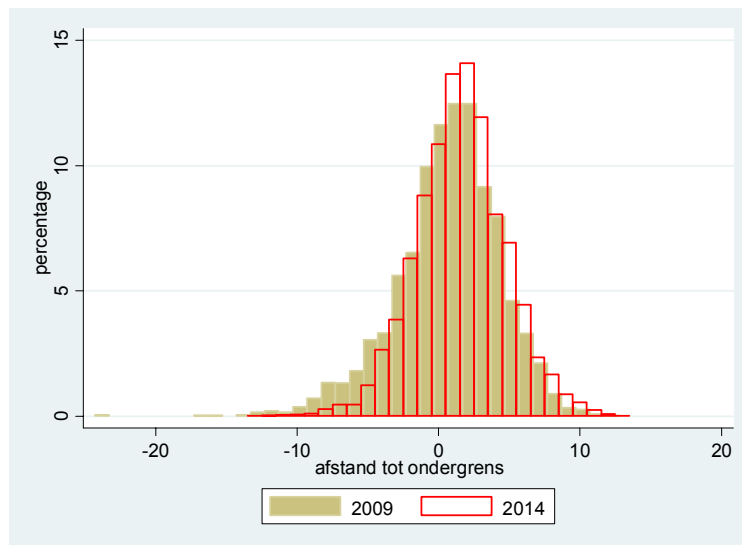
Figuur 5.7 laat de verdeling zijn van de scores van leerlingen op de Cito eindtoets in 2010 en 2013 (doordat de eindtoets in het voorjaar wordt afgenomen zijn dat zijn de schooljaren 2009/10 en 2012/13). De grafiek laat een aantal opmerkelijke sprongen zien. Deze ontstaan doordat het cito het aantal juiste vragen omzet naar deze citoscore. Het aantal ruwe scores dat samen wordt genomen in één citoscore is niet altijd gelijk. De grafiek laat drie verschuivingen zien. Ten eerste is het percentage leerlingen met een score van 550 in 2013 aanzienlijk groter dan in 2010. Ten tweede komen scores tussen 538 en 549 juist minder voor in 2013 in vergelijking met 2010, maar vaker scores tussen de 518 en 537.

Figuur 5.7 Verdeling van de score op de Cito eindtoets in 2010 en 2013



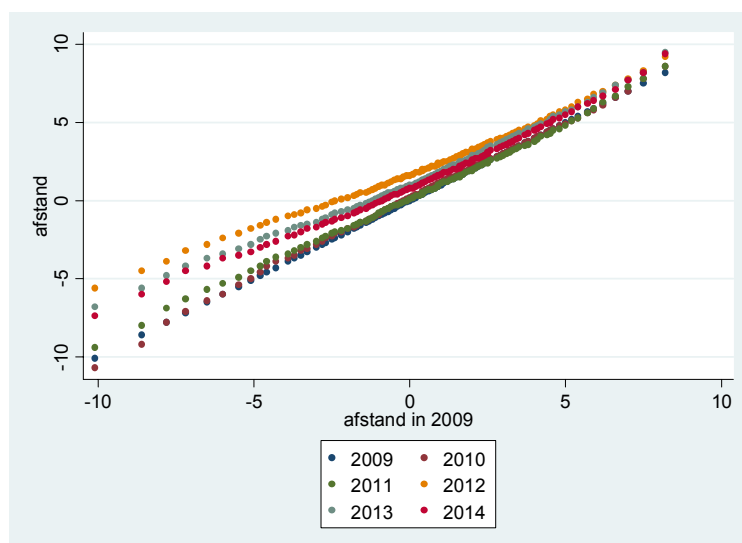
Voor de beoordeling van de resultaten van een school zijn niet de individuele scores van leerlingen van belang, maar de gemiddelden per school. De inspectie hanteert daarbij een ondergrens die afhangt van het percentage gewichtenleerlingen op de betreffende school. Voor de schoolgemiddeldes zijn langere tijdreeksen beschikbaar. Figuur 5.8 geeft de verdeling van het schoolgemiddelde van de score op de Cito eindtoets voor 2009 en 2014. Hierbij wordt gekeken naar de score ten opzichte van de genoemde ondergrens (de 0 in de grafiek). Omdat de gemiddelde scores van kleine scholen nogal fluctueren is hierbij alleen gekeken naar scholen waarbij minstens 20 leerlingen de eindtoets hebben gemaakt in het betreffende jaar.

Figuur 5.8 Verdeling van de afstand tot de ondergrens voor de gemiddelde score op de Cito eindtoets in 2009 en 2014.



In de eerste plaats is, naar verwachting, de verdeling van de gemiddelde scores veel compacter is dan de verdeling van individuele scores. In 2009 kwamen gemiddelde scores onder de door de inspectie vastgestelde ondergrens meer voor dan in 2014. De grootste toename doet zich voor bij gemiddelde scores rond de ondergrens. Figuur 5.9 laat deze ontwikkeling van gemiddelde toetsscores zien aan de hand van percentielen. Op de horizontale as staat de score van de percentielen in 2009. Op de verticale as de score van deze percentielen in latere jaren. De grafiek laat duidelijk zien dat er in de betreffende periode sprake is van een stijging van de score en dat deze groei zich voornamelijk heeft voorgedaan in het deel van de verdeling dat onder de ondergrens van de inspectie ligt.

Figuur 5.9 Verschuiving van de percentielen van de verdeling van de afstand tot de ondergrens van 2009 tot 2014.



Tabel 5.1 laat zien welk percentage van de scholen per jaar hoger scoort dan de ondergrens van de inspectie. Vooral in 2012 is dit percentage flinke toegenomen. Dat was het eerste jaar waarin de

indicator in de huidige vorm wordt gebruikt in het toezichtkader. De tweede kolom geeft aan hoe hoog dit percentage scholen dat de norm haalt zou zijn geweest, als alle scholen een gelijke toename van de gemiddelde score op de eindtoets ten opzichte van 2009 hadden gerealiseerd. Voor 2010 en 2011 ligt deze toename hoger dan de feitelijk toename, maar vanaf 2012 lager. Dat betekent dat vanaf 2012 met name scholen die dicht onder de ondergrens zaten betere resultaten hebben behaald.

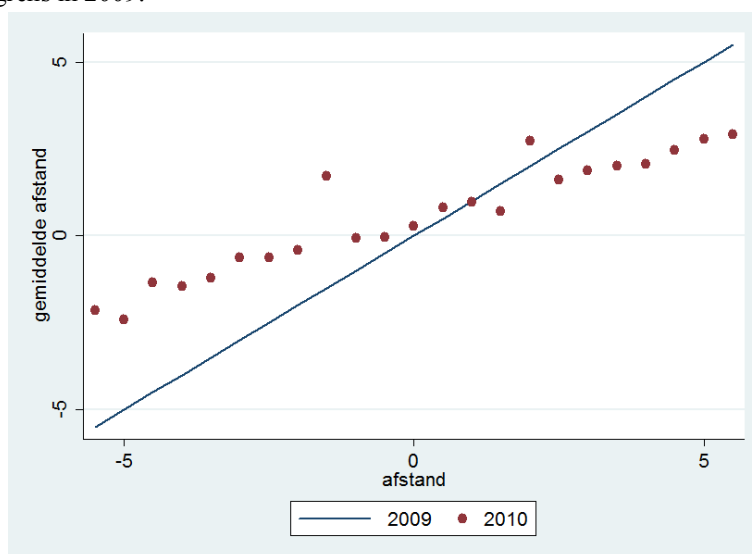
Tabel 5.1 Percentage scholen dat een gemiddelde cito eindtoets heeft boven de ondergrens en het percentage scholen dat een gemiddelde score boven de ondergrens zou hebben gehad als de stijging in de scores gelijk was geweest voor alle scholen.

Jaar	percentage boven ondergrens	percentage boven ondergrens bij uniforme stijging
2009	60.3	60.3
2010	61.5	63.9
2011	61.4	63.9
2012	81.1	76.7
2013	73.3	72.0
2014	70.8	70.3

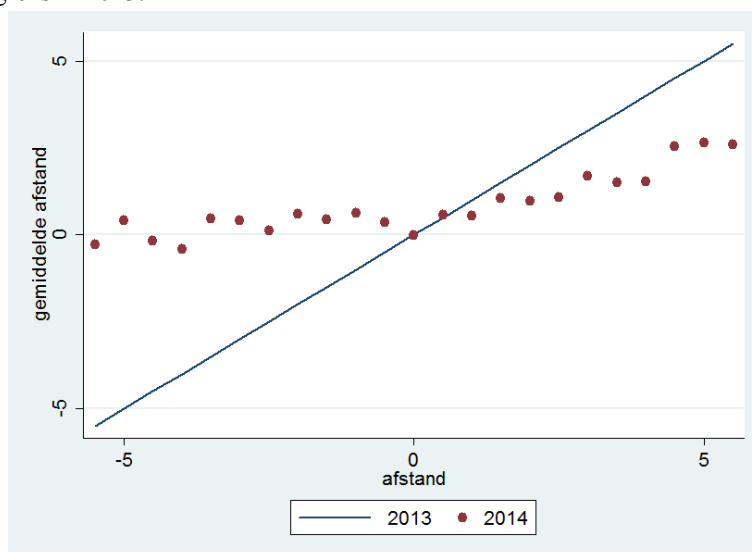
Het lijkt er dus op dat het beoordelingscriterium dat door de inspectie wordt gehanteerd een effect heeft gehad op de resultaten die scholen boeken en dat het met name de scholen zijn die onder de vastgestelde ondergrens lagen die beter zijn gaan scoren.

Als een school in een bepaald jaar onder de ondergrens van de inspectie scoort, kan het nog steeds aan de norm voldoen door in een van de twee andere jaren wel voldoende te scoren. Het is daarom interessant om de gemeten prestaties in een bepaald jaar te relateren aan de score die scholen in het voorafgaande jaar behaalden. De figuren 5.10 en 5.11 relateren de resultaten in 2010 en 2014 aan de scores die scholen gemiddeld haalden in respectievelijk 2009 en 2013. Ieder punt geeft aan welke afstand de score van een school had ten opzichte van de ondergrens in 2010 of 2014 gegeven de afstand van de score tot de ondergrens in het voorgaande jaar.

Figuur 5.10 Gemiddelde afstand van Cito-score tot ondergrens in 2010 gerelateerd aan de afstand tot de ondergrens in 2009.



Figuur 5.11 Gemiddelde afstand van Cito-score tot ondergrens in 2014 gerelateerd aan de afstand tot de ondergrens in 2013.



Wat opvalt is dat er in 2009/10 sprake is van een vrijwel lineair verband tussen de scores in beide jaren, terwijl er in dit patroon in 2013/14 een duidelijke knik in dit verband zit. De afvlakking in de score van 2010 ten opzichte van 2009 wordt veroorzaakt door de zogenaamde *regression to the mean*. Bij de scores in 2009 zitten uitschieters naar boven en beneden, in 2010 halen deze scholen weer scores die dichterbij hun trendmatige niveau liggen. De knik in 2013/14 is echter opmerkelijk. Deze impliceert dat scholen die lager dan de ondergrens scoren meer inspanning leveren voor een hogere score in 2014 dan scholen die al boven de ondergrens zitten.

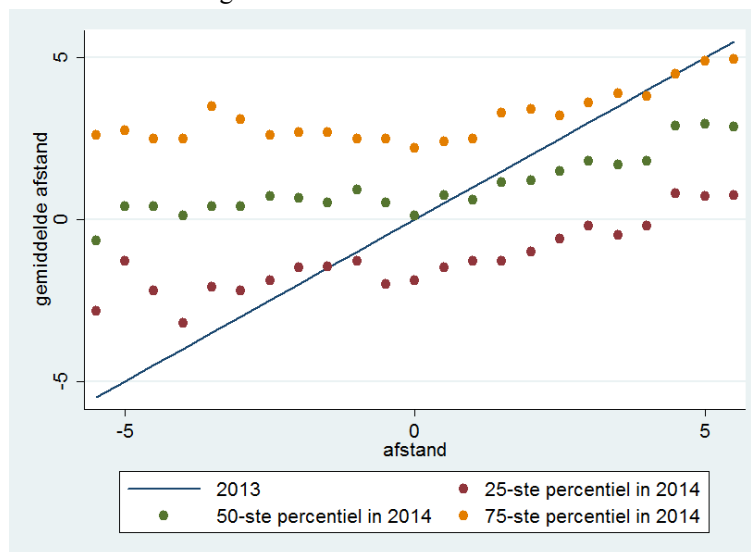
Tabel 5.2 geeft per jaar de coëfficiënten van de regressie van bovenstaande grafieken. De eerste coëfficiënt is de hellingshoek voor de relatie tussen de gemiddelde score van een school in het ene jaar, gerelateerd aan de score het jaar daarvoor voor de gevallen die score in het basisjaar groter was dan de ondergrens van de inspectie. De tweede coëfficiënt is het verschil van de deze hellingshoek voor gevallen waar de score lager was dan de ondergrens. In alle jaren behalve 2009 wordt er knik in het verband gevonden, die statistisch significant is.

Tabel 5.2 Hellingshoek van het verband tussen de afstand tot de ondergrens in een basisjaar en de gemiddelde afstand tot de ondergrens in het opvolgende jaar en de afwijking van dat verband voor scholen die in het basisjaar een Cito score hadden onder de ondergrens.

jaar	coëfficiënt	se	verschil coëfficiënt	se
2009	0.719	0.228	-0.184	0.442
2010	0.672	0.051	-0.403	0.101
2011	0.390	0.048	-0.316	0.096
2012	0.357	0.050	-0.244	0.124
2013	0.402	0.054	-0.248	0.116

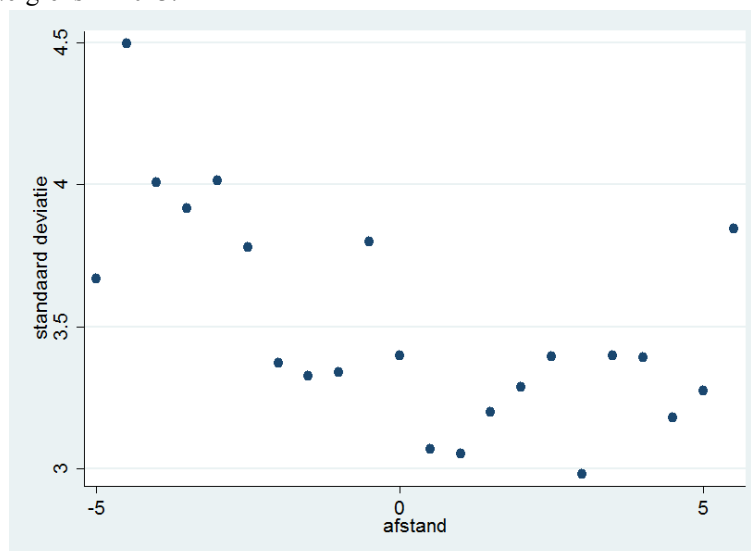
Figuur 5.12 geeft voor 2013/14 hetzelfde verband, maar nu met percentielen in plaats van gemiddelde scores. In de grafiek staan het 25-ste, 50-ste en 75-ste percentiel. Naarmate de school lager scoort in het basisjaar neemt de spreiding van de scores in het volgende jaar toe.

Figuur 5.12 25-ste, 50-ste en 75-ste percentiel van afstand van Cito score tot ondergrens in 2014 gerelateerd aan de afstand tot de ondergrens in 2013.



Figuur 5.13 geeft een beeld van de standaard deviatie van deze spreiding. Ook hier blijkt sprake te zijn van een knik in het verband rond de ondergrens. Met name bij scholen die onder de ondergrens presteren is er dus sprake van een verdere spreiding van de schoolresultaten in het volgende jaar. Dit lijkt er op te wijzen dat er verschillen tussen scholen zijn in de mate waarop ze reageren op een score beneden de ondergrens. Bij scholen boven de ondergrens doen deze verschillen zich niet voor.

Figuur 5.13 Standaarddeviatie van afstand van Cito score tot ondergrens in 2014 gerelateerd aan de afstand tot de ondergrens in 2013.



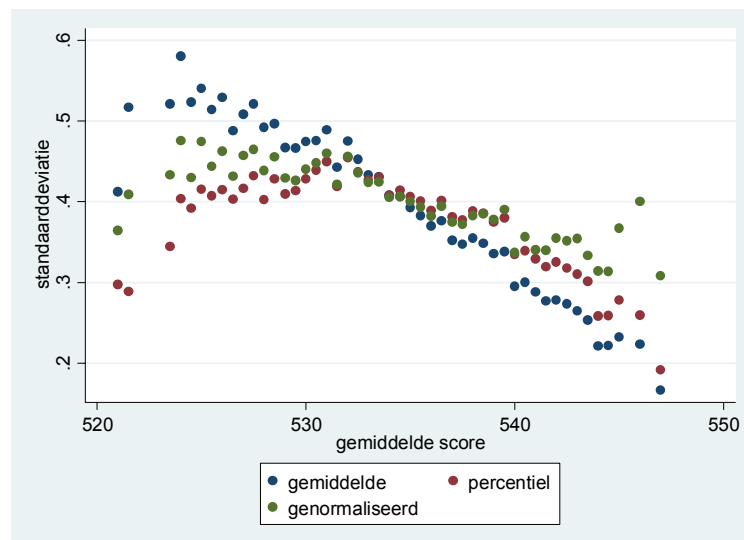
Vanwege de regel dat een school minimaal één keer in de drie jaar een gemiddelde score boven de ondergrens moet hebben, zou het denkbaar zijn dat hun strategie is om één jaar te pieken zodat daarmee de resultaten van een periode van drie jaar verzekerd kunnen worden. Van de 531 scholen die zowel in 2012 als 2013 een gemiddelde onder de ondergrens behaalden, heeft maar liefst 52 procent een score boven de ondergrens in 2014. Dit wekt wellicht de suggestie dat er sprake is van zo'n piekjaar. Toch blijkt dit niet het geval te zijn. De kans dat een school in één van de twee jaren

(bijvoorbeeld 2013 of 2014) de ondergrens haalt, blijkt niet groter te zijn dan verwacht zou worden op grond van de kans dat dit in één van die jaren wordt behaald.

De gemiddelde scores van scholen zijn aan grote schommelingen onderhevig. Deels komt dit doordat het aantal leerlingen in groep 8 niet heel groot is en daardoor het toeval een belangrijke rol speelt bij de uitslag, maar deel komt dit ook door de verdeling van de scores op de Cito eindtoets. Zoals figuur 5.7 laat zien heeft deze verdeling een lange staart aan de onderkant van de verdeling. De consequenties hiervan zijn dat de verschillen in scores tussen leerlingen aan de onderkant aanzienlijk groter zijn dan aan de bovenkant van de verdeling. Als men denkt aan een rangorde van alle leerlingen, dan betekent een verschuiving van 10 procent aan de onderkant van de verdeling een veel grotere toename van de Citoscore dan een verschuiving van 10 procent aan de bovenkant.

Figuur 5.14 laat zien hoe groot de standaarddeviatie van de scores van scholen is. Op de horizontale as staat de score die men zou behalen als er sprake zou zijn van zeer veel leerlingen. De simulatie is uitgevoerd voor scholen met minstens 20 leerlingen in groep 8. Er wordt een vergelijking gemaakt met de percentielscore en de genormaliseerde verdeling van de scores. Bij scholen met hoge gemiddelde scores loopt de standaard deviatie op ongeacht de methode die wordt gebruikt. Dit komt doordat scholen heterogener worden naarmate ze in dit segment een lagere score hebben. Bij de scholen met een lage score loopt de variantie het sterkst op voor de gemiddelde Citoscore en is dat minder het geval bij de andere maten. Het toevallig wel of niet een leerling hebben die zeer laag scoort heeft een grote invloed op het gebruik van de gemiddelde Citoscore als maat voor de prestatie van scholen. Dit geeft hen een prikkel om zich met name op deze zwakste leerlingen te richten, en maakt hen ook zeer gevoelig voor het toevallig in school hebben van een zeer zwakke leerlingen.

Figuur 5.14 Standaarddeviatie van gemiddelde score van een school, gerelateerd aan het gemiddelde niveau van die school voor drie verschillende vormen van de indicator (gemiddelde, percentiel, genormaliseerde score).

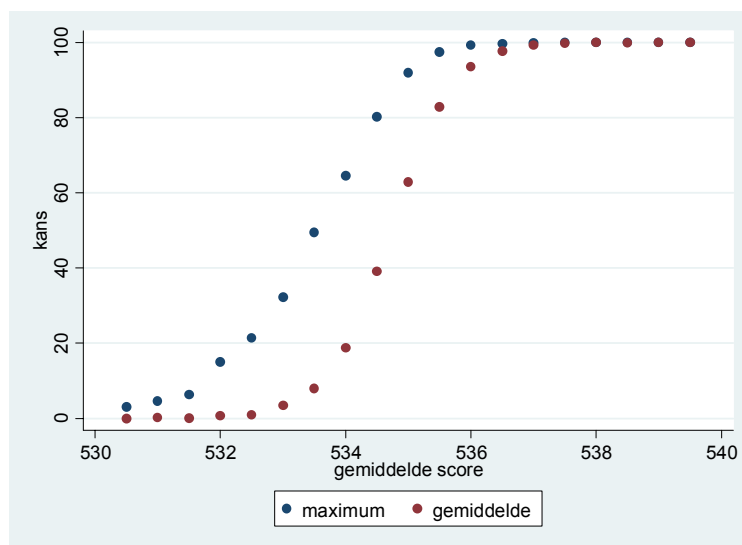


Ook de regel dat scholen één van de drie jaar een gemiddelde boven de ondergrens moeten behalen, maakt de invloed van toeval bij de berekening groter. In figuur 5.15 wordt getoond hoe groot de kans is dat een school de driejarennorm haalt bij de huidige regel en bij de fictieve regel waarin naar het gemiddelde over drie jaar wordt gekeken. Omdat het gemiddelde een strengere norm is, zijn de kansen

daar lager, maar van belang is dat de curve sneller stijgt. Het grijze gebied tussen vrijwel zeker de norm halen, en vrijwel zeker de norm niet halen is groter bij het gebruik van de regel dat een keer de ondergrens moet worden behaald.

In lijn met deze bevinding dat de invloed van hogere scores aan de onderkant veel groter is dan de invloed van hogere scores aan de bovenkant, vinden wij ook dat scholen die in een jaar de ondergrens niet behalen, in het volgende jaar vooral hoger scoren bij de leerlingen met het laagste resultaten. Dat kan betekenen dat men (strategisch) meer energie in deze leerlingen steekt omdat de invloed van hogere scores bij hen ook de grootste invloed heeft op het gemiddelde, maar het kan ook komen doordat extra inspanningen bij leerlingen die al beter scoren minder snel tot uiting komt in de toetsresultaten.

Figuur 5.15 Kans dat een school hoger scoort dan 535, gerelateerd aan het gemiddelde niveau van die school voor het geval de maximale of de gemiddelde score in drie jaar maatgevend is.



6 Conclusies en aanbevelingen

In dit onderzoek is gekeken naar de invloed die opbrengstindicatoren in het toezichtkader van de inspectie kunnen hebben op het gedrag van basisscholen. De inspectie beoordeelt het functioneren van scholen met indicatoren die staan geformuleerd in het toezichtkader. Omdat een goede beoordeling van belang is voor scholen levert dat een prikkel op om op die indicatoren voldoende te scoren. Vooral voor scholen dicht bij de grenswaarde van zo'n indicator zal de invloed naar verwachting het grootst zijn. In het ideale geval leidt het gebruik van dergelijke indicatoren tot een verbeterde onderwijskwaliteit. Het is echter ook mogelijk dat de indicatoren scholen er toe aanzetten om gedrag te vertonen dat weliswaar gunstig uitwerkt voor de indicatoren, maar geen of zelfs een negatieve invloed heeft op de kwaliteit. Omdat indicatoren altijd globaal van karakter zijn, kan het ook zijn dat het toezichtkader de scholen belemmert om dingen te doen die eigenlijk goed zouden zijn voor het onderwijs. Bij het gebruik van dergelijke indicatoren speelt dus altijd een trade-off van enerzijds de stimulerende werking en anderzijds de negatieve bijwerkingen.

Er zijn een aantal gesprekken gevoerd in het veld (PO-raad, AVS, directeuren en leerkrachten) om een algemeen beeld van de situatie te krijgen. In deze gesprekken komt naar voren dat men over het algemeen het idee steunt dat het goed is om met opbrengstindicatoren de kwaliteit van het onderwijs te stimuleren. Op basis van deze gesprekken is een vragenlijst opgesteld voor directeuren en leerkrachten. 79 procent van de ondervraagden vindt het goed of heel goed dat de inspectie met dergelijke opbrengstindicatoren werkt. Over de invulling van de indicatoren is men echter minder positief. In de gesprekken komt regelmatig het beperkende karakter van de indicatoren aan de orde. Slechts 57 procent van de ondervraagden denkt dat de gebruikte indicatoren een positieve effect op het onderwijs hebben.

Om de vraag te beantwoorden in hoeverre opbrengstindicatoren scholen mogelijk kunnen aanzetten tot strategisch gedrag of inperken in hun beleid, wordt in dit onderzoek een empirische analyse uitgevoerd. Uit de gegevens zoals die door de inspectie worden gebruikt zal op schoolniveau niet snel duidelijk worden dat er sprake is van strategisch gedrag. Dit komt doordat de uitkomsten ook het resultaat kunnen zijn van gewoon – niet strategisch – handelen en bovendien in hoge mate door toeval bepaald kunnen zijn en doordat de inspectie geen informatie heeft over de achterliggende redenen waarom scholen iets doen. Bepaalde gedragingen hoeven geen reactie te zijn op het toezichtkader, maar kunnen ook andere oorzaken hebben. Om toch een beeld te krijgen van de omvang van strategisch gedrag, moeten gegevens gebruikt worden die (door de inspectie) niet herleidbaar zijn tot de betreffende scholen. In dit onderzoek laten we zien in hoeverre geaggregeerde verdelingen een beeld geven van mogelijk strategisch gedrag dat heeft plaatsgevonden. Vaak zijn er op geaggregeerd niveau sporen te constateren van strategisch gedrag zonder dat dat aan een bepaalde school is toe te wijzen. Daarnaast zijn directeuren en leerkrachten anoniem ondervraagd waardoor eerlijke antwoorden niet tot repercussies hoeven te leiden. Omdat toch altijd de mogelijkheid bestaat dat een respondent niet graag volledige openheid van zaken geeft, hebben we ook gevraagd hoe men denkt dat zaken lopen bij andere scholen. Een groot verschil tussen het geaggregeerde antwoord over de eigen school en andere scholen suggereert dat dit het geval is.

Daarnaast zijn er een aantal vragen aan ouders voorgelegd over hoe zij het onderwijs percipiëren. Een discrepantie tussen de kijk van de ouders en die van de school zou er op kunnen wijzen dat de scholen in de vragenlijst een mooier beeld proberen te geven. Dat verschil vinden we echter niet terug.

In een analyse met grootschalige microbestanden is gekeken naar twee facetten van strategisch gedrag. De doorstroom van leerlingen en de resultaten op de Cito eindtoets. Hoewel het toezichtkader van de inspectie indicatoren bevat over de vertraging van leerlingen, spelen deze nauwelijks een rol in de feitelijke beoordeling door de inspectie. Ondanks dat geven toch veel directeuren en docenten aan dat ze rekening houden met de indicatoren van de inspectie over de doorstroom bij hun beslissingen. Uit de analyse komt naar voren dat er opvallend weinig scholen zijn met een zeer laag percentage onvertraagde doorstroom. Als de kans op vertraging alleen afhangt van kenmerken van de leerlingen, zou dit percentage puur vanwege toeval regelmatig zeer laag moeten zijn. De analyse laat zien dat dergelijke toevallige lage percentages opvallend weinig voorkomen. Kennelijk beslissen scholen die door toeval een jaargang leerlingen hebben die het vrij goed doen, sneller om individuele leerlingen langer over de basisschool te laten doen dan anders. Interessant is dat dit scholen betreft die over het algemeen ruim beneden de inspectienorm scoren met betrekking tot onvertraagde doorstroom. Bij een groot percentage scholen is sprake van vertraagde doorstroom die ver boven deze norm zit. Maar juist bij hen lijkt er geen sprake te zijn van strategisch gedrag.

Een tweede indicator die is geanalyseerd in grootschalige microbestanden is de Cito eindtoets. De inspectie hanteert een ondergrens voor de gemiddelde Cito eindtoets die afhangt van het percentage gewichtenleerlingen. Sinds 2009 is er een stijging in de scores bij de scholen die gemiddeld genomen onder deze ondergrens scoren. Bij de beter scorende scholen is er vrijwel geen verandering. Het lijkt er dus op dat de indicator inderdaad invloed heeft gehad op de scholen die potentieel geraakt worden door dit criterium. Of dit komt doordat de kwaliteit van het onderwijs is verbeterd, of doordat scholen via strategisch gedrag hun scores hebben weten te verhogen zonder dat dit feitelijk berust op beter onderwijs, kan niet uit deze data worden afgeleid. Scholen moeten over een periode van drie jaar eenmaal aan de betreffende ondergrens voldoen. Hoewel scholen die een grote kans hebben om onder deze ondergrens te scoren zich extra lijken in te spannen om de norm te halen, blijkt het niet zo te zijn dat men eens in de drie jaar éénmalig de norm weet te halen. Net als de indicatoren over vertraging is de gemiddelde toetsscore voor een aanzienlijk deel afhankelijk van toevalsfactoren.

Uit de enquête onder leerkrachten en directeuren komt overigens wel naar voren dat gedrag waarbij een positieve relatie is met de resultaten van de school zeker niet ongebruikelijk is. De grens tussen geoorloofd en niet geoorloofd gedrag is daarbij gradueel. Uit de enquête komen de volgende cijfers naar voren:

- 26 procent van de respondenten geeft aan bij het doorverwijzen naar het speciaal onderwijs op zijn minst een beetje rekening te houden met het toezichtkader van de inspectie.
- 31 procent doet dat bij beslissingen over de doorstroom van leerlingen.
- 40 procent van de respondenten geeft aan op zijn minst een beetje rekening te houden met het toezichtkader bij het buiten de eindtoets laten van leerlingen.
- 48 procent leest wel eens vragen van de eindtoets voor.
- 58 procent helpt op zijn minst leerlingen een beetje op weg bij de eindtoets.
- 69 procent bespreekt vooraf de lesstof van de vragen van de toets.
- 75 procent geeft leerlingen wel eens extra tijd bij het maken van de eindtoets.

- 82 procent stemt op zijn minst een beetje de lesstof af op de toetsen van het leerlingvolgsysteem.
- 85 procent houdt bij de aanschaf van lesmethodes rekening met de eindtoets basisonderwijs en
- 94 procent van de respondenten geeft aan dat leerlingen op hun school op zijn minst een beetje gerust worden gesteld bij het maken van de eindtoets.

Vooral bij de vragen rond de afname van de eindtoets blijkt er vaak een samenhang te zijn tussen de gemiddelde Cito eindtoetsresultaten van de school, en de ondergrens van de inspectie en de mate waarin de bevraagde gedragingen voorkomen. Bij sommige aspecten valt het grote verschil op in de mate waarin respondenten aangeven dat dit voorkomt op de eigen school en andere scholen. Het gaat hier om het niet aannemen van zwakke leerlingen vanwege de consequenties voor het beoordelingskader en het buiten de toets laten van leerlingen, het op weg helpen van leerlingen die moeite hebben met vragen en het voorlezen tijdens de toets. Een verschil in de aandacht voor taal en rekenen in de periode voor de eindtoets versus de periode na de eindtoets lijkt heel erg mee te vallen op grond van zowel de antwoorden van de leerkrachten en directeurs als de ouders. Leerkrachten geven vaker aan dan schoolleiders dat de opbrengstindicatoren een vrij sterke of zeer sterke rol spelen bij beslissingen over toelating, doorstroom en doorverwijzing, en zeggen ook vaker dat het op andere scholen voorkomt. Opmerkelijk is overigens dat leerkrachten en schoolleiders aangeven dat de indicator over doorstroom een rol speelt bij hun beleid, terwijl de inspectie aangeeft dat deze indicator niet of nauwelijks een rol speelt bij het toezicht. Kennelijk betekent het feit dat de indicator is opgenomen in het toezichtkader, dat sommige scholen er van uitgaan dat de indicator er wel toe doet.

Aanbevelingen

Het beeld dat ontstaat over mogelijk strategisch gedrag van scholen is dat er een duidelijk waarneembaar effect van het toezichtkader is op de scores van de scholen op de eindtoets, maar dat dit effect in omvang bescheiden is. Ook bij de vertraagde doorstroom is strategisch gedrag aanwezig, maar ook daar is de omvang bescheiden. Bovendien lijkt dit strategisch gedrag geen reactie op de betreffende opbrengstindicator te zijn. De mate waarin scholen hun gedrag lijken aan te passen aan het toezichtkader, met name de gemiddelde score op de eindtoets, is echter opvallend groot. Uit de enquête komt naar voren dat scholen dus veel doen om hun door de indicatoren gemeten resultaten een klein beetje te verbeteren. Twee oorzaken hiervan zijn, dat bij de beoordeling van de scholen erg veel afhangt van de gemiddelde eindtoets en dat de invloed van toeval bij deze score vrij groot is. Dat leidt tot een gevoel bij scholen overgeleverd te zijn aan een beoordeling die niet goed recht doet aan hun feitelijke prestaties. In het jargon van management control: De indicatoren zijn niet congruent en nauwkeurig.

De mate waarin scholen hun aanpak aanpassen om betere toetsscores te behalen is opvallend hoog. Een mogelijke reactie zou dan ook zijn om de regels over de afname van de eindtoets aan te scherpen om daarmee de eerlijkheid van de afnamecondities te bevorderen. Dergelijke aanpassingen nemen echter niet de druk weg die scholen ervaren, die voor een groot deel de oorzaak zijn van het gedrag van deze scholen. Het ligt daarom in onze ogen meer voor de hand om de beoordeling minder sterk afhankelijk te maken van de eindtoetsscore en te gaan werken met een bredere set indicatoren. Daarnaast is het van belang om te erkennen dat indicatoren slechts een ruwe benadering zijn van wat er speelt op scholen, en in het proces dan ook meer ruimte te laten voor een dialoog met de scholen

over de achtergrond van de behaalde resultaten. Uit de gesprekken met schoolleiders kwam naar voren dat hieraan behoefte is, maar dat dit in sommige gevallen al wel wordt ervaren. Een mogelijkheid om dit te realiseren zou zijn, om naast een ondergrens voor de eindtoetscijfers waaronder de opbrengsten als onvoldoende worden gezien ook een ondergrens te hanteren voor een verder gesprek over de resultaten en de achtergrond daarvan. Bij scholen die resultaten behalen in die tussenzone wordt op basis van dat gesprek bepaald of de resultaten als onvoldoende worden gezien. De lagere ondergrens zou daarbij lager moeten worden dan de huidige ondergrens, de hogere ondergrens juist hoger. Daardoor ontstaat er een zone waarin de opbrengsten van een school niet alleen door de score op de eindtoets worden beoordeeld, maar op basis van een gesprek meer context in de beslissing kan worden betrokken. Voor een dergelijke dialoog lijkt het ook van belang dat duidelijkheid bestaat over welke indicatoren een rol spelen in het toezicht en welke niet. De huidige situatie dat er indicatoren in het toezichtkader staan die geen rol spelen, lijkt tot onduidelijkheid te leiden.

Daarnaast kunnen er maatregelen genomen worden om de gebruikte indicatoren minder gevoelig te maken voor toevallige fluctuaties. Ten eerste kan gekeken worden of de gemiddelde toetsresultaten over drie jaar boven een bepaalde ondergrens liggen, in plaats van de regel dat de resultaten minstens een keer in de drie jaar boven een ondergrens moeten liggen. Ten tweede zou de toetscore herschaald kunnen worden zodat de gemiddelde score minder gevoelig wordt voor toevallige uitschieters aan de onderkant. Bij dergelijke aanpassingen moet uiteraard de ondergrens opnieuw worden vastgesteld.

Beoordeling van scholen aan de hand van indicatoren heeft – ook in de ogen van de betrokkenen op de scholen – duidelijk een stimulerende werking voor de onderwijskwaliteit, maar kent ook ongewenste neveneffecten. Dit is inherent aan dit instrument. Het is dus van belang om te zoeken naar een goede balans tussen beide. Bij die optimale balans zal het altijd het geval zijn dat soms op basis van indicatoren conclusies worden getrokken die niet helemaal recht doen aan de werkelijkheid van de school, en zal er voldoende ruimte voor scholen moeten blijven om zelf invulling te geven aan hun koers. Als gestreefd wordt om volledig te voorkomen dat beoordelingen van de inspectie niet helemaal recht doen aan de situatie op de school, zal het instrument zijn kracht verliezen. Als omgekeerd geprobeerd wordt alle vormen van in de ogen van de inspectie ongewenst gedrag van scholen in te perken met aanvullende regels en voorschriften, zal daarmee ook de speelruimte voor scholen om te doen wat goed is voor de leerlingen verdwijnen. Er ontstaat dan een situatie waarin scholen bezig zijn met indicatoren in plaats van met leerlingen.

In zo'n goede balans ontstaat dus een situatie waarin de inspectie op basis van de indicatoren waarop zij scholen beoordeelt, geen volledig beeld krijgt van hoe het er voor staat in het onderwijs. De inspectie zal dus altijd moeten werken met een bepaalde mate van onzekerheid. Zowel voor de monitoring van de kwaliteit van het Nederlandse onderwijs als zodanig, als voor de zoektocht naar de juiste balans tussen inperking en vrijheid, is het van belang om buiten de indicatoren die gebruikt worden bij het toezicht, gegevens te verzamelen over het onderwijs en over de mate van strategische gedrag waar de scholen niet op afgerekend kunnen worden. Dit onderzoek laat zien dat zowel door de analyse van geaggregeerde verdelingen als door anonieme ondervraging relevante informatie verzameld kan worden hiervoor. Gezien de zorg om verenging van het onderwijsaanbod als reactie op het grote belang van toetsresultaten op het gebied van taal en rekenen, lijkt het van belang om voor een bredere set uitkomsten gegevens te verzamelen om in ieder geval op stelselniveau de ontwikkeling van het onderwijs te kunnen volgen. Andere uitkomsten dan taal en rekenen zijn vaak moeilijk te meten op

een wijze die niet makkelijk beïnvloed kan worden door de school. Daardoor ligt het gebruik van indicatoren voor dergelijke uitkomsten niet voor de hand. Door aanvullend gegevens te verzamelen op stelselniveau kan wel gemonitord worden in hoeverre de druk op de toets deze neveneffecten heeft. Iets vergelijkbaars geldt voor de metingen van taal- en rekenvaardigheden zelf. Als duidelijk is dat toetsresultaten op scholen in sterke mate onder druk staan vanwege de beoordeling door de inspectie, dan is het van belang om de opbrengsten op het gebied van taal en rekenen ook te monitoren met instrumenten die niet gebruikt worden bij de beoordeling van scholen. Een instrument als PPO is hier heel geschikt voor. Door aan de ene kant overlap te creëren tussen de vragen die in de eindtoets worden opgenomen en de vragen die in PPO zitten, maar anderzijds ook juist een breder spectrum van taal- en rekenvaardigheden te toetsen, kan bekeken worden in hoeverre de vaardigheden die uit de eindtoets blijken ook buiten die toets om aanwezig blijken te zijn.

Bijlage A Kernkader primair onderwijs

A. Opbrengsten	
1	De opbrengsten liggen op het niveau dat op grond van de kenmerken van de leerlingenpopulatie verwacht mag worden.
1.1*	De resultaten van de leerlingen aan het eind van de basisschool liggen ten minste op het niveau dat op grond van de kenmerken van de leerlingenpopulatie mag worden verwacht.
1.2*	De resultaten van de leerlingen voor Nederlandse taal en voor rekenen en wiskunde tijdens de schoolperiode liggen ten minste op het niveau dat op grond van de kenmerken van de leerlingenpopulatie mag worden verwacht.
1.3	De leerlingen doorlopen in beginsel de school binnen de verwachte periode van 8 jaar.
1.4	Leerlingen met specifieke onderwijsbehoeften ontwikkelen zich naar hun mogelijkheden.
1.5	De sociale competenties van de leerlingen liggen op een niveau dat mag worden verwacht.
1.6	De adviezen van de leerlingen voor het vervolgonderwijs zijn in overeenstemming met de verwachtingen op grond van de kenmerken van de leerlingenpopulatie.
1.7	De leerlingen functioneren naar verwachting in het vervolgonderwijs.
B. Onderwijsleerproces	
2	De aangeboden leerstofinhouden bereiden de leerlingen voor op vervolgonderwijs en samenleving.
2.1*	Bij de aangeboden leerinhouden voor Nederlandse taal en rekenen en wiskunde betreft de school alle kerndoelen als te bereiken doelstellingen.
2.2*	De leerinhouden voor Nederlandse taal en voor rekenen en wiskunde worden aan voldoende leerlingen aangeboden tot en met het niveau van leerjaar 8.
2.3	De leerinhouden in de verschillende leerjaren sluiten op elkaar aan.
2.4*	De school met een substantieel aantal leerlingen met een leerlinggewicht biedt bij Nederlandse taal leerinhouden aan die passen bij de onderwijsbehoeften van leerlingen met een taalachterstand.
2.5	De school heeft een specifiek aanbod om sociale competenties te ontwikkelen.
2.6	De school heeft een aanbod gericht op bevordering van actief burgerschap en sociale integratie, met inbegrip van het overdragen van kennis over en kennismaking met de diversiteit in de samenleving.
3	De leraren geven de leerlingen voldoende tijd om zich het leerstofaanbod eigen te maken.
3.1	De leraren maken efficiënt gebruik van de geplande onderwijstijd.
4	Het schoolklimaat wordt gekenmerkt door veiligheid en respectvolle omgangsvormen.
4.1	De ouders zijn betrokken bij de school door de activiteiten die de school daartoe onderneemt.
4.2	De leerlingen voelen zich aantoonbaar veilig op school.
4.3	Het personeel voelt zich aantoonbaar veilig op school.
4.4	De school heeft inzicht in de veiligheidsbeleving van leerlingen en personeel en in de incidenten die zich op het gebied van sociale veiligheid op de school voordoen.
4.5	De school heeft een veiligheidsbeleid gericht op het voorkomen van incidenten in en om de school.
4.6	De school heeft een veiligheidsbeleid gericht op de afhandeling van incidenten in en om de school.
4.7	Het personeel van de school zorgt ervoor dat de leerlingen op een respectvolle manier met elkaar en anderen omgaan.
5	De leraren leggen duidelijk uit, organiseren de onderwijsactiviteit efficiënt en houden de leerlingen taakbetrokken.
5.1*	De leraren geven duidelijke uitleg van de leerstof.
5.2*	De leraren realiseren een taakgerichte werksfeer.
5.3*	De leerlingen zijn actief betrokken bij de onderwijsactiviteiten.
6	De leraren stemmen aanbod, instructie, verwerking en onderwijstijd af op verschillen in ontwikkeling tussen de leerlingen.
6.1	De leraren stemmen de aangeboden leerinhouden af op verschillen in ontwikkeling tussen de leerlingen.

	6.2	De leraren stemmen de instructie af op verschillen in ontwikkeling tussen de leerlingen.
	6.3	De leraren stemmen de verwerkingsopdrachten af op verschillen in ontwikkeling tussen de leerlingen.
	6.4	De leraren stemmen de onderwijstijd af op verschillen in ontwikkeling tussen de leerlingen.
C. Zorg en begeleiding		
7a		De leraren volgen systematisch de vorderingen van de leerlingen.
	7.1*	De school gebruikt een samenhangend systeem van genormeerde instrumenten en procedures voor het volgen van de prestaties en de ontwikkeling van de leerlingen.
	7.2	De leraren volgen en analyseren systematisch de voortgang in de ontwikkeling van de leerlingen.
7b		Specifiek voor SBO: De school begeleidt de leerlingen zodat zij zich naar hun mogelijkheden ontwikkelen.
	S7.3	De school stelt bij plaatsing voor iedere leerling een ontwikkelingsperspectief vast.
	S7.4	De school volgt of de leerlingen zich ontwikkelen conform het ontwikkelingsperspectief en maakt naar aanleiding hiervan beredeneerde keuzes.
8		De leerlingen die dat nodig blijken te hebben, krijgen extra zorg.
	8.1	De school signaleert vroegtijdig welke leerlingen zorg nodig hebben.
	8.2	Op basis van een analyse van de verzamelde gegevens bepaalt de school de aard van de zorg voor de zorgleerlingen.
	8.3	De school voert de zorg planmatig uit.
	8.4	De school evalueert regelmatig de effecten van de zorg.
	8.5	De school zoekt de structurele samenwerking met ketenpartners waar noodzakelijke interventies op leerlingniveau haar eigen kerntaak overschrijden.
D. Kwaliteitszorg		
9		De school heeft een systeem voor kwaliteitszorg.
	9.1	De school heeft inzicht in de onderwijsbehoeften van haar leerlingenpopulatie.
	9.2	De school evalueert jaarlijks de resultaten van de leerlingen.
	9.3	De school evalueert regelmatig het onderwijsleerproces.
	9.4	De school werkt planmatig aan verbeteractiviteiten.
	9.5	De school borgt de kwaliteit van het onderwijsleerproces.
	9.6	De school verantwoordt zich aan belanghebbenden over de gerealiseerde onderwijskwaliteit.
	9.7	De school draagt zorg voor de kwaliteit van het onderwijs gericht op bevordering van actief burgerschap en sociale integratie, met inbegrip van het overdragen van kennis over en kennismaking met de diversiteit van de samenleving.
E. Wet- en regelgeving		
N1a		De inspectie is in het bezit gesteld van de geldende schoolgids (art. 16, lid 2 en 3, WPO).
N1b		In de schoolgids is informatie opgenomen over standaard gecontroleerde onderdelen (art. 13 WPO).
N2a		De inspectie is in het bezit gesteld van het geldende schoolplan (art. 16, lid 1 en 3, WPO).
N2b		In het schoolplan is informatie opgenomen over een of meer standaard gecontroleerde onderdelen (art. 12 WPO).
N3a		De inspectie is in het bezit gesteld van het geldende zorgplan (art. 19, lid 3, WPO).
N3b		In het zorgplan is informatie opgenomen over een of meer standaard gecontroleerde onderdelen (art. 19, lid 2, WPO).
N4a		Er is voldoende onderwijstijd geprogrammeerd om te kunnen voldoen aan de daarvoor gestelde minima (art. 8, lid 7 onder b, WPO).
N4b		Er is maximaal zevenmaal een onvolledige schoolweek geprogrammeerd voor de groepen 3 t/m 8 (artikel 8, lid 7 onder b, WPO).

Noot: De indicatoren met een * zijn de normindicatoren die een belangrijke rol spelen bij de vaststelling van het toezichtarrangement.

Bijlage B Gespreksleidraad voor gesprekken met schoolleiders en leerkrachten

In het toezicht van de inspectie zijn opbrengstindicatoren de afgelopen tijd een belangrijkere rol gaan spelen. In dit gesprek zouden we met u willen bespreken hoe u dat als school ervaart. Vindt u dat de indicatoren die de inspectie gebruikt een goed beeld opleveren van wat er op uw school gebeurt? Hoe ervaart u deze vorm van beoordeling? Beïnvloedt het toezichtkader van de inspectie de manier waarop les wordt gegeven en de beleidsmatige keuzes die worden gemaakt?

We stellen u eerst algemene vragen over hoe u de invloed van opbrengstindicatoren ervaart. Daarna gaan we in op een aantal specifieke zaken die bij deze indicatoren een rol spelen. Daarna komen weer terug op uw kijk op het gebruik van indicatoren door de inspectie.

- 1) Bent u goed op de hoogte van het toezichtkader dat de inspectie hanteert?
- 2) Vindt u het – in het algemeen – een goede zaak dat de inspectie met indicatoren de kwaliteit van scholen in kaart probeert te brengen?
- 3) Vindt u de indicatoren zoals die worden gehanteerd door de inspectie de juiste indicatoren?
- 4) Welke indicatoren zijn in uw ogen niet goed? En waarom niet?
- 5) Welke indicatoren ontbreken in uw ogen? En waarom?
- 6) Kreeg uw school meer aanmeldingen van nieuwe leerlingen dan dat er plaatsen beschikbaar zijn?
 - a. Indien ja, welke criteria gebruikte de school om leerlingen te selecteren?
 - b. Indien nee, welke criteria zou de school gebruiken om leerlingen te selecteren als er meer aanmeldingen zijn dan plaatsen?
- 7) Speelt op uw school de beoordeling van de inspectie een rol bij deze selectiecriteria?
- 8) In hoeverre denkt u dat op andere scholen de beoordeling van de inspectie een rol bij deze selectiecriteria?
- 9) Welk percentage van de leerlingen op uw school blijft zitten ergens gedurende de schoolperiode of blijft een jaar langer in de kleuterklas?
- 10) Welk percentage van de leerlingen op andere scholen, denkt u, blijft zitten ergens gedurende de schoolperiode of blijft een jaar langer in de kleuterklas?
- 11) Speelt op uw school de beoordeling van de inspectie of de bekostiging een rol bij de beslissing om leerlingen een jaar over te laten doen?
- 12) Indien ja, zou u als de inspectie deze invloed niet had graag vaker, of juist minder vaak leerlingen een jaar over willen laten doen?
- 13) Denkt u dat op andere scholen de beoordeling van de inspectie of de bekostiging een rol bij de beslissing om leerlingen een jaar over te laten doen?
- 14) Welke toets wordt op uw school in groep 8 afgenomen?
- 15) Wat is de reden voor uw school om deze toets af te nemen?
- 16) Speelt voor uw school bij de keuze voor deze toets de beoordeling door de inspectie een rol?
- 17) Denk u dat bij andere scholen bij de keuze voor deze toets de beoordeling door de inspectie een rol speelt?
- 18) Hoe vaak oefenen op uw school leerlingen in groep 8 met het maken van de eindtoets?
- 19) Hoe vaak denkt u dat leerlingen in groep 8 op andere scholen oefenen met het maken van de eindtoets?
- 20) Speelt op uw school bij de manier waarop u leerlingen voorbereid op de eindtoets de beoordeling door de inspectie een rol?
- 21) Hoe?

- 22) Denkt u dat op andere scholen bij de manier waarop zij hun leerlingen voorbereiden op de eindtoets de beoordeling door de inspectie een rol speelt?
- 23) Hoe?
- 24) Speelt op uw school bij de manier waarop u de eindtoets afneemt de beoordeling door de inspectie een rol?
- 25) Hoe?
- 26) Denkt u dat op andere scholen de manier waarop zij de eindtoets afnemen de beoordeling door de inspectie een rol speelt?
- 27) Hoe?
- 28) Hoe hoog denkt u dat het percentage leerlingen is dat niet deelneemt aan de eindtoets op uw school?
- 29) Hoe hoog denkt u dat gemiddeld het percentage leerlingen is dat niet deelneemt aan de eindtoets op andere scholen?
- 30) Wat zijn op uw school de voornaamste redenen waarom een leerling niet meedoet aan de eindtoets?
- 31) Speelt op uw school bij de beslissing om een leerling niet deel te laten nemen aan de eindtoets de beoordeling van de inspectie een rol?
- 32) Denkt u dat op andere scholen de beslissing om een leerling niet deel te laten nemen aan de eindtoets de beoordeling van de inspectie een rol speelt?
- 33) Hoe hoog is op uw school het percentage leerlingen is dat de eindtoets nog eens over doet of dat waarbij u een andere eindtoets afneemt?
- 34) Hoe hoog denk u dat op andere scholen het percentage leerlingen is dat de eindtoets nog eens over doet of dat waarbij u een andere eindtoets afneemt?
- 35) Wat zijn op uw school de voornaamste redenen om bij leerlingen de eindtoets nog eens over te laten doen of een andere eindtoets af te nemen?
- 36) Speelt op uw school de beoordeling van de inspectie een rol om bij leerlingen de eindtoets nog eens over te laten doen of een andere eindtoets af te nemen?
- 37) Denkt u dat op andere scholen de beoordeling van de inspectie om bij leerlingen de eindtoets nog eens over te laten doen of een andere eindtoets af te nemen een rol speelt?
- 38) Hoe vaak komt het op uw school voor dat een leerling gedurende de schoolperiode wordt doorverwezen naar het speciaal onderwijs?
- 39) Hoe vaak denkt u dat het op andere scholen voorkomt dat een leerling gedurende de schoolperiode wordt doorverwezen naar het speciaal onderwijs?
- 40) Speelt op uw school bij het al dan niet doorverwijzen van leerlingen naar het speciaal onderwijs de beoordeling door de inspectie of de bekostiging van uw school een rol?
- 41) Denkt u dat op andere scholen bij het al dan niet doorverwijzen van leerlingen naar het speciaal onderwijs, de beoordeling door de inspectie of de bekostiging van uw school een rol speelt?
- 42) Als de beoordeling van de inspectie of de bekostiging geen rol zou spelen, zou u dan vaker of juist minder vaak leerlingen naar het speciaal onderwijs doorverwijzen?

Tot slot van dit interview willen we nog kijken wat uw algemene oordeel is over het gebruik van indicatoren door de inspectie.

- 43) Vindt u dat het gebruik van indicatoren door de inspectie een stimulerend effect heeft voor het onderwijs op deze school of vindt u het effect er van juist negatief?
- 44) Vindt u dat de inspectie het gebruik van indicatoren aan zou moeten passen? Hoe dan?
- 45) Als u één aspect van het toezichtkader zou mogen veranderen, wat zou dat dan zijn?

Bijlage C Vragenlijst ouders

Groep 2

In de reguliere oudervragenlijst van het jaarlijkse onderzoek in het kader van de Limburgse Onderwijsmonitor zitten een aantal vragen voor ouders van leerlingen in groep 2 die nuttig zijn voor dit onderzoek.

Schoolkeuze

1. Uit hoeveel basisscholen kon u kiezen?
 - Ik kon niet kiezen
 - Ik kon kiezen uit twee
 - Ik kon kiezen uit drie of meer
 - Weet ik niet meer
- 2a. Was deze school uw eerste keuze?
 - Ja
 - Nee
- 2b. Als deze school niet uw eerste keuze was
Welke school was uw eerste keuze _____?
Waarom is uw kind niet naar de school van uw eerste keuze gegaan _____?

Groep 8

Aan de reguliere oudervragenlijst van het jaarlijkse onderzoek in het kader van de Limburgse Onderwijsmonitor zijn een aantal vragen toegevoegd voor ouders van leerlingen in groep 8.

Advies voortgezet onderwijs

1. Welk advies voor Voortgezet Onderwijs kreeg het kind?
(bij combinaties: alles aanvinken wat van toepassing is)
 - VMBO-praktisch of beroepsgericht
 - VMBO (algemeen)
 - VMBO-theoretische leerweg (MAVO)
 - HAVO
 - HAVO-VWO
 - VWO
 - (nog) geen advies
 - anders, namelijk.....
2. Was u tevreden met dat schooladvies?
 - zeer ontevreden
 - ontevreden
 - geen oordeel
 - tevreden
 - zeer tevreden

Cito eindtoets

3. Heeft uw kind deelgenomen aan de Cito eindtoets?
 - ja
 - nee
4. Als uw kind niet heeft deelgenomen, wat was de reden hiervoor?
5. In hoeverre zijn de volgende stellingen van toepassing?
(zeker niet / niet / twijfel / wel / zeker wel)(in sommige gevallen alleen niet/wel)
 - Op school is veel geoefend voor de Cito eindtoets
 - Mijn kind was erg gespannen voor de Cito eindtoets
 - Tijdens het afnemen van de Cito eindtoets was de sfeer op school prettig
 - In groep 8 is in de periode *tot februari* veel nadruk gelegd op taal en rekenen
 - In groep 8 lag in de periode *vanaf februari tot nu* veel nadruk op taal en rekenen
 - Bij het vervolgdadvies krijgen de ouders voldoende inspraak
 - Door leesproblemen is het maken van een toets moeilijk voor mijn kind
 - Mijn kind heeft officieel de diagnose dyslectisch
 - De school raadde ons ooit aan ons kind een jaar over te laten doen
 - Wat ons betreft had ons kind beter een jaar langer over de basisschool kunnen doen
 - Mijn kind zou tijdens de basisschool mogelijk de overstap maken naar speciaal onderwijs

Leerling met speciale behoeften

6. Is uw kind een "zorgleerling" vanwege leer- of gedragsmoeilijkheden of handicap?
 - Ja
 - Nee
 - Weet niet

Zittenblijven

7. Heeft uw kind ooit een jaar over gedaan?
 - Nee
 - Ja, tussen groep 2 en 3
 - Ja, in een hogere groep

Bijlage D Vragenlijst schoolleiders en leerkrachten

Vragen voor de enquête voor schoolleiders en leerkrachten

Inleiding 1

(voor de eerste persoon die je aan de lijn krijgt, bijv. conciërge of secretaresse)

U spreekt met ... van ITS Radboud Universiteit. We voeren momenteel een onderzoek uit naar hoe scholen het toezicht door de Inspectie van het Onderwijs ervaren. Daarvoor voeren we korte telefonische gesprekken door het hele land om de mening van schoolleiders en leerkrachten te peilen. Zou ik kunnen spreken met de directeur [of: de leerkracht van groep X] om hem/haar enkele vragen hierover te stellen? Het duurt ongeveer vijf minuten en de resultaten worden anoniem verwerkt.

Inleiding 2

(voor de directeur of leerkracht die je gaat interviewen)

U spreekt met ... van ITS Radboud Universiteit. We voeren momenteel een onderzoek uit naar hoe scholen het toezicht door de Inspectie van het Onderwijs ervaren. Daarvoor voeren we korte telefonische gesprekken door het hele land om de mening van schoolleiders en leerkrachten te peilen. Zou ik u enkele korte vragen over dit onderwerp mogen stellen?

Evt. meer uitgebreide toelichting

In het toezicht van de Inspectie van het Onderwijs zijn de opbrengstindicatoren van de school van groot belang. Met opbrengstindicatoren bedoelen we onder andere de resultaten van de eindtoets, indicatoren over de ontwikkeling van leerlingen met een specifieke onderwijsbehoefte en dergelijke. In deze korte telefonische enquête willen wij u vragen hoe u dit als schoolleider/leerkracht ervaart. We willen graag weten wat u vindt van deze vorm van beoordeling, en of het toezicht van de inspectie invloed heeft op de keuzes die u of die de school maakt, bijvoorbeeld bij het onderwijs geven en in het beleid. De enquête duurt ongeveer vijf minuten en de resultaten worden anoniem verwerkt. Ik stel u steeds een vraag en dan geef ik de antwoordcategorieën waaruit u kunt kiezen.

- a. Wat is uw functie?
 - Directeur
 - Leerkracht
- b. Indien directeur: geeft u naast uw functie als directeur ook nog les?
 - Ja
 - Nee
- c. Indien leerkracht of “ja” bij vraag b: In welke groep of groepen geeft u les? (meerdere antwoorden mogelijk)
 - Groep 3
 - Groep 4
 - Groep 5
 - Groep 6
 - Groep 7
 - Groep 8

Tabel D1 Algemeen overzicht wie welke vragenblokken krijgt.

Schoolleider	Leerkracht groep 3-7	Leerkracht groep 8
Algemeen	Algemeen	Algemeen
Toelating		
Zittenblijven	Zittenblijven	
	Lvs toets oefenen	
Eindtoets algemeen		Eindtoets algemeen
Keuze lesmethode	Keuze lesmethode	Keuze lesmethode
Doorverwijzing SBO	Doorverwijzing SBO	Doorverwijzing SBO
		Eindtoets specifiek

Noot: Als de directeur ook lesgeeft, krijgt hij ook de vragen voor die betreffende leerkracht.

Algemeen

1. Vindt u zelf dat u goed op de hoogte bent van de inhoud van het toezichtkader en de indicatoren van de Inspectie van het Onderwijs?
 - Heel goed
 - Goed
 - Niet zo goed
 - Helemaal niet goed
 (als antwoord op vraag 1 is “helemaal niet”, dan vraag 2 overslaan)
2. Vindt u dat de opbrengstindicatoren die worden gehanteerd door de Inspectie van het Onderwijs de juiste indicatoren zijn?

Evt. toelichting: Met opbrengstindicatoren bedoelen we dan onder andere de resultaten van de eindtoets, resultaten van voortgangstoetsen, indicatoren over de ontwikkeling van leerlingen met specifieke onderwijsbehoefte, en dergelijke

 - Ja, helemaal
 - Grotendeels
 - Enigszins
 - Helemaal niet
 - Geen mening
3. Vindt u het – in het algemeen – een goede zaak dat de inspectie met opbrengstindicatoren werkt om de kwaliteit van scholen in kaart probeert te brengen?
 - Heel goed
 - Goed
 - Slecht
 - Heel slecht
 - Geen mening
4. Vindt u dat het gebruik van opbrengstindicatoren door de inspectie een positief effect heeft op het onderwijs op de school of vindt u het effect er van juist negatief?
 - Heel positief
 - Een beetje positief
 - Vrij negatief
 - Negatief
 - Weet ik niet

Toelating

5. Stel dat er een nieuwe leerling wordt aangemeld voor groep 1 van wie wordt vermoedt dat het een zwakke leerling zal zijn. In hoeverre wegen bij de beslissing om deze leerling al dan niet toe te laten de consequenties voor de opbrengstbeoordeling door de Inspectie van het Onderwijs mee?
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Weet ik niet
6. In hoeverre denkt u dat dit op scholen in uw omgeving gebeurt? Dus, dat bij de toelating van leerlingen van wie het vermoeden bestaat dat ze zwak zijn rekening wordt gehouden met de impact die dit heeft op de beoordeling door de Inspectie van het Onderwijs?
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Weet ik niet

Zittenblijven

7. Als er op uw school beslist moet worden of een leerling doorgaat naar een hogere groep of een jaar overdoet, in hoeverre speelt daarbij dan de beoordeling door de inspectie een rol?
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Weet ik niet
8. In hoeverre denkt u dat op scholen in uw omgeving de beoordeling door de inspectie een rol speelt bij de beslissing om leerlingen een jaar over te laten doen?
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Weet ik niet

Lvs toetsen

9. Als u weet dat een bepaald onderdeel van een geplande toets uit het leerlingvolgsysteem op dat moment in de lesmethode nog niet voldoende aan bod is geweest, in hoeverre besteedt u daar dan extra aandacht aan in de les?
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Nvt geen voortgangstoetsen
10. In hoeverre denkt u dat dit op scholen in uw omgeving gebeurt? Dus, buiten de lesmethode om extra aandacht geven aan onderwerpen die in de leerlingvolgsysteem-toets aan bod komen.
 - Helemaal niet
 - Een beetje
 - Vrij sterk

- Heel sterk
- Weet ik niet

Eindtoets algemeen

11. Hoe vaak, schat u, maken leerlingen in groep 8 op uw school een oefentoets ter voorbereiding van de eindtoets? Kunt u het aantal keer aangeven?
 - Getal:
 - Weet ik niet
12. Hoe vaak, schat u, maken leerlingen in groep 8 op scholen in uw omgeving een oefentoets ter voorbereiding van de eindtoets? Kunt u het geschatte aantal keer aangeven?
 - Getal:
 - Weet ik niet
13. In hoeverre speelt bij de keuze om een leerling wel of niet mee te laten doen aan de eindtoets de beoordeling door de Inspectie van het Onderwijs een rol?
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - nvt
14. In hoeverre denkt u dat dit op scholen in uw omgeving meespeelt? Dus, dat ze leerlingen niet mee laten doen aan de eindtoets vanwege de beoordeling door de Inspectie van het Onderwijs?
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk

Keuze lesmethode

15. In hoeverre speelt bij de keuze voor de aanschaf van lesmethoden een rol dat deze aansluiten op de eindtoets en de toetsen uit het leerlingvolgsysteem?
Evt. toelichting: Zodat leerlingen niet een toets krijgen met zaken die niet in de lesmethode aan bod zijn geweest.
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Nvt, geen voortgangstoetsen
16. In hoeverre denkt u dat dit op scholen in uw omgeving gebeurt? Dus, dat scholen zich bij de aanschaf van een lesmethode laten leiden door het feit dat deze goed aansluit op de eindtoets of leerlingvolgsysteem-toetsen?
 - Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Weet ik niet

Doorverwijzen naar speciaal onderwijs

17. In hoeverre speelt op uw school bij het al dan niet doorverwijzen van leerlingen naar het speciaal onderwijs de beoordeling van de Inspectie van het Onderwijs een rol?
- Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Weet ik niet
18. In hoeverre denkt u dat dit op scholen in uw omgeving meespeelt?
- Helemaal niet
 - Een beetje
 - Vrij sterk
 - Heel sterk
 - Weet ik niet
19. Als de schoolresultaten geen rol zouden spelen bij de beoordeling van de Inspectie van het Onderwijs zou u dan vaker of juist minder vaak leerlingen naar het speciaal onderwijs doorverwijzen?
- Vaker
 - Hetzelfde als nu
 - Minder vaak

Eindtoets specifiek

In hoeverre zijn de volgende stellingen van toepassing?

(zeker niet / niet / twijfel / wel / zeker wel / niet van toepassing)

20. De kinderen waren erg gespannen voor de eindtoets
21. In groep 8 is in de periode *tot februari* veel nadruk gelegd op taal en rekenen
22. In groep 8 lag in de periode *vanaf februari tot nu* veel nadruk op taal en rekenen

Hoe vaak denkt u dat het volgende gedrag bij de afname van de eindtoets voorkomt op uw school?

(Nooit / bijna nooit / soms/ regelmatig / vaak)

23. Vooraf lesstof bespreken die aan de orde gaat komen in de eindtoets
24. Tijdens de eindtoets, leerlingen die moeite hebben met een vraag op weg helpen
25. Tijdens de eindtoets, leerlingen geruststellen
26. Tijdens de eindtoets, leerlingen extra tijd geven
27. Tijdens de eindtoets vragen voorlezen

Hoe vaak denkt u dat het volgende gedrag bij de afname van de eindtoets voorkomt op scholen in uw omgeving?

(Nooit / bijna nooit / soms/ regelmatig / vaak)

28. Vooraf lesstof bespreken die aan de orde gaat komen in de eindtoets
29. Tijdens de eindtoets, leerlingen die moeite hebben met een vraag op weg helpen
30. Tijdens de eindtoets, leerlingen geruststellen
31. Tijdens de eindtoets, leerlingen extra tijd geven
32. Tijdens de eindtoets vragen voorlezen